

THIS WEEK

EDITORIALS

WORLD VIEW The new model army needed to tackle financial crises **p.143**

ANTS Social insects take one for the diseased team **p.144**



MISSILE North Korea prepares controversial rocket launch **p.146**

Buyer beware

An investigation by Nature shows the scale of the market for unapproved stem-cell therapies in China. Hype and unrealistic hope must not be allowed to undermine genuine promise.

On the Internet, you can find advertisements for stem-cell remedies for every kind of disease or injury. Companies also promise that the cells will improve appearance or provide a ‘rejuvenating’ energy boost. The message — that stem-cell therapies need some work, but are an accepted part of medicine — is as clear as it is wrong.

But repeat this mantra enough — as it is repeated endlessly online — and the promises can start to seem real. In some places, they certainly look real. As we reveal on page 149, these ‘cures’ are offered in real clinics in China, where real nurses and doctors inject people with stem cells in various formulations from various sources — apparently convinced that they are helping patients. It looks and feels routine.

China has tried to crack down on unapproved treatments and it is not the only place where patients can buy these therapies: stem-cell companies also take advantage of gaps in regulation enforcement in the United States (see *Nature* **483**, 13–14; 2012). But in China the problem is more widespread.

Promoters of such unproven and unapproved ‘treatments’ liken stem-cell therapy to other once-revolutionary therapies, such as organ transplantation. Doctors confess that they can’t guarantee that the stem cells will work, but they do guarantee that the procedures are safe. If they weren’t, say advocates, we would hear about it. So why not try?

This circular logic makes the apparent infiltration of stem-cell technologies into the medical mainstream even more worrisome. The more willing patients and medics are to believe, the less they look for true clinical data, and the less doctors are forced to produce it.

Compare the emergence of stem-cell therapies with the introduction of psychosurgery. Like stem-cell practitioners today, doctors in the 1930s and 1940s felt that the need for lobotomy was urgent enough

to bypass the requirement for clinical evidence. Results were reviewed selectively, with pacifying brain damage sometimes taken as a stabilizing ‘cure’ for schizophrenia or nervous disorders. One promoter even shared the 1949 Nobel Prize in Physiology or Medicine for his part in developing the procedure. There was no long-term follow-up. In the end, doctors mutilated the brains of thousands of patients over the course of decades before critics were able to cast enough doubt on lobotomy to

“Acceptance is already overtaking clinical evidence, with no systematic follow-up.”

halt its use. The widespread acceptance made it difficult for people to realize that these procedures were actually doing harm: you don’t see a problem if you don’t bother looking for it.

Of course, there is much legitimate research into stem cells, including many controlled clinical trials. It would be a shame if they were tainted by association with historical failures.

However, judging from *Nature*’s investigation in China, acceptance is already overtaking clinical evidence, with no attempt at systematic follow-up of treatments. If stem-cell therapies result in cancer or immunological disease, no one will know.

This does not stop people from outside China flocking to the country to take advantage of the stem-cell therapies offered there and promoted online with glowing endorsements. The clinics are certainly set up to make foreigners feel at home. Set aside from the teeming Chinese hospitals, stem-cell treatment centres have orderly nurse stations, well-lit rooms and good bedside care. What is lacking is controlled clinical trials, reliable data and government approval. If the dedicated medical workers at the clinics don’t see the problems, they need to look harder. If they really want to help their patients, they should seek to prove that the treatments work, rather than just assuming that they do. ■

Honest work

The plagiarism police deserve thanks for defending the honour of the PhD.

Last week, Hungary’s President Pál Schmitt was forced to resign because of plagiarism detected in his 1992 PhD thesis on physical education. Tivadar Tulassay, rector of Budapest’s prestigious Semmelweis University, showed admirable courage by standing up to the Hungarian establishment to revoke the thesis a few days earlier, after experts appointed by the university declared that Schmitt’s thesis “failed to meet scientific and ethical standards”. Tulassay, a cardiovascular researcher, has since assumed personal responsibility for his university’s decision to revoke Schmitt’s title.

The affair has remarkable parallels with that of Germany’s former defence minister, Karl-Theodor zu Guttenberg, who resigned in March last year after his own PhD thesis, in law, had been revoked by the University of Bayreuth.

Like Schmitt, zu Guttenberg tried at first to deny plagiarism charges, then to underplay them, and he enjoyed powerful political support — until protests by a movement of honest PhD holders made his situation untenable. Plagiarism hunters have other prominent personalities in their sights, and are not necessarily going to be stopped just because a thesis is not in electronic form — if suspicion is high, they will digitize it themselves.

In many central European countries, an academic title is a decided advantage for a political career; clearly, some ambitious politicians think nothing of obtaining such a title by cheating. We can thank the plagiarism hunters — whatever their individual motives — for exposing dishonesty among those who govern us, and for defending the honour of a PhD. The only safe doctorate these days is an honestly acquired one. ■

JEFF SWENSEN



Scientists and bankers — a new model army

Bankers must now surrender more information on their activities. Scientists should use it to build better system-wide financial models, says John Liechty.

The financial system is in a credit-confidence trap. Like a badly balanced ship trying not to capsize during a storm, banks and financiers are unwilling to make loans or accept collateral in exchange for securing debts — they fear being overwhelmed by the next wave of crisis. Even though the first sovereign default, in Greece, has passed, there is no obvious port in this storm.

How did the system get so far out of balance? And how can scientists — accustomed to modelling complex events to explain and make predictions — help to create a financial system that is more self-stabilizing? Existing financial models failed to predict the crisis of 2008 and the follow-on crisis of 2011–12. They missed the huge system-wide risks that developed as banks promoted an undisciplined supply of mortgages and created an increasingly complex web of relationships through legal contracts that transferred risk throughout the financial markets.

Financial bonds based on these mortgages (and other assets) were seen as risk-free and cheap, and banks used them as both capital and collateral. But when house prices plummeted, the bonds were useless for securing even short-term loans for the banks, which suddenly faced cash shortages. The banks held assets that were potentially worthless, and they were all interconnected. If one firm went down, everyone else was vulnerable.

Market forces function only if all risks are fairly priced. The system-wide risks that these bonds held were not taken into account, so the bonds were sold too cheaply. A clear scientific goal, therefore, is to build better system-wide models of the global financial system. Both the industry and regulators could use such models to judge financial risk and make decisions.

True, scientists are not blameless with regard to the recent collapse. They helped to create the models that the banks routinely used to measure risk. But those models lacked crucial data — on common holdings and trading behaviours, for instance, and on the interconnections between firms and the capacity of markets to execute trades.

For commercial reasons, banks have historically been reluctant to share this kind of information, but that is changing. Legislation in the United States now allows regulators to collect such data from banks, pension funds, insurance companies and other big players in the financial markets. Regulators in Europe are following suit, and hopefully Asia will as well. As a result, we will soon be able to model and identify potential system-wide risks.

To get an idea of the challenge of modelling system-wide financial risk, consider an indoor shopping mall that charges people to enter and exit. To model the movement of shoppers, we could build a purely statistical model of the door

traffic, and in most situations this would be sufficient. However, in extreme situations, such as a fire, the system would change dramatically. Shoppers would rush for the nearest exits and ticket-takers would get overwhelmed and close their doors. Then shoppers would rush to the next set of doors. In such cases, a statistical model would get it horribly wrong.

Equilibrium or structural models of the same system would track and predict the motives — and, therefore, the movement — of each individual. In normal times, these models are too complex and hard to calibrate — imagine trying to quantify all of the reasons that people go shopping. But in times of distress, as the shoppers' motives and behaviour converge (*get out!*), the model output improves.

To fully understand and predict the dynamics of a market in crisis, we have to understand the capacity of the market-makers (the ticket-takers at the door of the shopping mall) and the demand for assets when prices lurch significantly away from present levels (the number of shoppers trying to get out versus the number trying to get in). The new data will allow models to do this for the first time.

Clearly, regulators have a responsibility to build such models and to use them to monitor for potential crisis. To do this, they will need to leverage expertise among scientists by supporting and encouraging research in universities and labs, and by hosting the more applied work to maintain confidentiality. Bankers should join this effort too, if only to avoid forcing regulators to use crude tools to set prices on these risks — through capital ratios or transaction taxes, for example.

Bankers should work in parallel and form an industry group that collects system-wide data from its members, organizes resources for scientists to develop the necessary models, and creates a secure and confidential infrastructure for members to determine the price of system-wide risks. The industry already has a group that does something similar — the international Operational Riskdata eXchange Association, which shares operational-loss data among member firms.

Everyone would benefit if bankers were to engage with scientists to build the infrastructure needed to price system-wide risk. Banks could get feedback about common holdings and trading strategies, which would allow them to adjust their behaviour and avoid following the herd. Regulators would have extra market information to help them to determine when to act to ensure stability. And the rest of us could have increased confidence in the financial system. ■

**A CLEAR SCIENTIFIC
GOAL IS TO BUILD
BETTER
MODELS
OF THE GLOBAL
FINANCIAL
SYSTEM.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/mzcyl6

John Liechty is director of the Center for the Study of Global Financial Stability and Professor of Marketing and Statistics at Pennsylvania State University in University Park.
e-mail: jcl12@psu.edu

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

ECOLOGY

A look at backyard biodiversity

The choices city dwellers make when deciding which plants to cultivate in their gardens or yards could affect the function and health of the wider plant ecosystem.

Sonja Knapp at the Helmholtz Centre for Environmental Research in Halle, Germany, and her group compared the characteristics and diversity of plants in 137 residential yards with those of a nearby nature reserve. They found that yard plant species were more closely related to each other, shorter-lived, faster-growing and more likely to be self-pollinating.

As yard plants spread into natural habitats, the ability of those ecosystems to respond to environmental change could be reduced, the authors say.

Ecology <http://dx.doi.org/10.1890/11-0392.1> (2012)

ANIMAL BEHAVIOUR

Licking ants fight fungal infection

Healthy ants that rub up against infected counterparts or even lick pathogenic fungal spores off them may be immunizing themselves and, ultimately, protecting their whole colony.

Sylvia Cremer at the Institute of Science and Technology Austria in Klosterneuburg and her colleagues infected ants

(*Lasius neglectus*; pictured) with fluorescently labelled fungal spores (*Metarhizium anisopliae*) and released them among healthy members of their colony. The authors

found that spores frequently transferred to healthy ants, resulting in low-level infection. Genetic analysis revealed that these minor infections upregulated a set of immune-system genes that bolstered the ants' anti-fungal defences. Computer modelling suggests that this 'social immunization' actively stimulates the ants' immune systems, allowing the colony as a whole to fight infection. **PLoS Biol.** 10, e1001300 (2012)



M. KONRAD/IST AUSTRIA



MARKA/SUPERSTOCK

GEOLOGY

Hot tuff not so tough

Tuffs are volcanic rocks commonly used as building materials despite their notorious weakness — and at least one popular tuff could pose an even greater hazard in the event of a fire.

Michael Heap at the University of Strasbourg in France and his colleagues examined three types of tuff commonly used in buildings in the Neapolitan region of Italy (pictured). Two exhibited no reduction in strength after thermal stressing, but the most commonly used one,

known as Neapolitan yellow tuff, lost 80% of its compressive strength as temperatures reached 1,000 °C. This is explained by the fact that Neapolitan yellow tuff contains zeolite minerals that are sensitive to heat.

The team suggests that the results be considered in establishing regional fire codes and recommends similar tests for building tuff in other regions.

Geology 40, 311–314 (2012)

MOLECULAR BIOLOGY

MicroRNAs boost gene variation

Small RNA molecules that regulate and stabilize the expression of certain genes in humans may also promote and preserve variations in gene expression between individuals and ethnic groups.

Jian Lu and Andrew Clark at Cornell University in Ithaca, New York, examined the expression profiles of protein-coding genes that are influenced by microRNAs (miRNAs) and were obtained from multiple human populations. The authors compared these profiles with

those of genes not targeted by miRNAs. Expression of some of the miRNA-regulated genes varied little across populations, or between humans, chimpanzees and macaques. However, most differed from one individual to another and between ethnic groups.

Genome Res. <http://dx.doi.org/10.1101/gr.132514.111> (2012)

NANOTECHNOLOGY

Lasers sort particles by size

Gold nanoparticles have a range of biomedical uses, from detecting tumours to delivering drugs. However, their size is important because

it affects their optical and mechanical properties, as well as their toxicity. Martin Ploschner and his colleagues at the University of St Andrews, UK, report an efficient way to sort gold nanoparticles by size using laser light.

The team aimed green and red lasers at a thin layer of water containing a mixture of gold nanoparticles 100 and 130 nanometres in diameter. The green light's frequency matched that of the electrons in the smaller nanoparticles. This resonance enhanced forces acting on the particles, pushing them in one direction. The red light interacted with the larger particles, moving them in the opposite direction.

The researchers suggest that the method could sort nanoparticles more finely than current methods, which rely on centrifugation.

Nano Lett. <http://dx.doi.org/10.1021/nl204378r> (2012)

GENETICS

Fewer imprinted genes at re-count

Most mammalian cells have one maternal and one paternal copy of most genes, but some genes carry a molecular signature or 'imprint' that silences one copy. Tomas Babak at Stanford University in California and his team mapped the imprinted genes in mouse brains and found far fewer than recent estimates had suggested.

In 2010, two studies found more than 1,300 imprinted genes in the mouse brain, ten times more than traditional counts. The increase was attributed to improved RNA-sequencing technology. When Babak *et al.* repeated the experiments, they found only 13% of the imprinted genes first identified by the 2010 studies and uncovered statistical weaknesses that resulted in many false-positive signals. Using a different analytical approach, the authors identified roughly 50 new candidate imprinted genes.

Having a catalogue of imprinted genes is important for understanding why imprinting occurs and how it can go awry.

PLoS Genet. **8**, e1002600 (2012)

MOLECULAR IMAGING

Follow the lymph vessels

Lymph vessels grow as wounds heal and cancers spread — a process that can be imaged in living animals, researchers demonstrate in mice.

Lymph vessels often sprout at sites of inflammation, and their growth has been linked to tumour metastasis. Sagrario Ortega at the Spanish National Cancer Research Centre in Madrid and her colleagues genetically engineered a mouse to express a luminescent protein under the control of the gene *Vegfr3*, a lymphatic marker.

The team imaged live mice, tracking vessel growth during embryo development, wound healing and inflammation. They also watched as lymph vessels grew at the edge of melanoma tumours and in lymph nodes infiltrated by the cancer. This vessel growth may aid the spread of cancer to distant organs, the authors say. *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1115542109> (2012)

MATERIALS

A graphene window on liquids

By using graphene membranes as viewing 'windows', researchers have filmed nanocrystals growing in liquids at atomic resolution.

Studying structures in liquids at the atomic level is challenging because the imaging technique of choice, transmission electron microscopy, requires that samples be in a vacuum to maximize their interactions with the electron beam. Air-tight capsules can be used to enclose liquids, but are thick and made of materials

COMMUNITY CHOICE

The most viewed papers in science

CANCER

Small, cancer-resistant mice

HIGHLY READ
on www.cell.com
in March

Boosting the levels of a tumour-suppressor protein in mice makes them smaller and more metabolically efficient, as well as resistant to cancer.

Pier Paolo Pandolfi at Beth Israel Deaconess Medical Center in Boston, Massachusetts, and his colleagues genetically engineered mice to have additional copies of *Pten*, a gene that is mutated or deleted in many cancers. The mice are smaller than normal because they have fewer cells. When injected with a carcinogen, the animals developed tumours later than controls.

The transgenic mice burn energy at a higher rate. Cells from these mice consume less glucose than normal mouse cells but generate more ATP — the energy molecule created during cellular respiration — indicating a more efficient metabolism.

Increasing levels of the PTEN protein could offer a therapeutic approach to preventing both cancer and obesity. *Cell* **149**, 49–62 (2012)

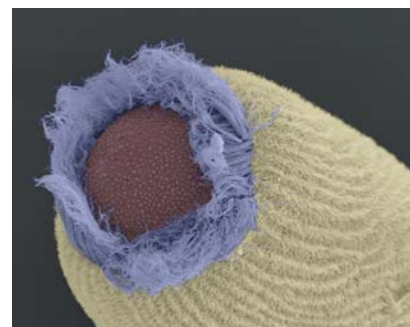
that interfere with passing electrons, resulting in a blurred picture. Membranes made of graphene — atomically thin sheets of carbon atoms — are both impermeable to liquids and much more transparent to electrons.

Paul Alivisatos at the Lawrence Berkeley National Laboratory in Berkeley, California, and his colleagues used these graphene windows to create atomic-resolution movies of platinum nanocrystals clumping together in a liquid. *Science* **336**, 61–64 (2012)

ZOOLOGY

Vision with no nervous system

Sponge larvae can detect light despite lacking a nervous system or the photosensitive 'opsin' proteins found in all other known animal eyes. Instead, another pigment called cryptochrome may underlie the light-sensing ability of the sponge *Amphimedon queenslandica* (pictured), report Todd Oakley at the University of California, Santa Barbara, and his colleagues. Cryptochromes mainly



absorb blue light and, in other animals, have been implicated in functions from setting circadian rhythms to sensing magnetic fields. The authors identified two cryptochrome genes in the sponge. One, *Aq-Cry2*, is expressed in the 'ring eyes' of *A. queenslandica* larvae and has an absorbance peak similar to the wavelengths that trigger larval activity.

Because eye evolution in other animals has always involved opsins, the use of cryptochrome represents a separate lineage of eye evolution, the team suggests. *J. Exp. Biol.* **215**, 1278–1286 (2012)

► **NATURE.COM**

For the latest research published by Nature visit:
www.nature.com/latestresearch

S. LEYS

SEVEN DAYS

The news in brief

POLICY

North Korea launch

North Korea has continued its plans to launch a long-range rocket in defiance of international pressure not to break a ban on missile testing. With a launch planned for between 12 and 16 April, the rocket (which North Korea says carries a weather satellite) was moved into position at the Sohae Satellite Launching Station on the country's northwestern coast last week. The plans have brought a swift end to February's agreement with the United States for food aid in return for a moratorium on nuclear testing, uranium enrichment and long-range ballistic missile development (see *Nature* 483, 128; 2012).

Astrophysics review

NASA's exoplanet-hunting space telescope, Kepler, is among nine astrophysics missions to have their lifetimes extended after a performance review by external scientists. Both Kepler and Swift, which detects γ -ray bursts, will run through to 2016. Although it recommended extensions for all the missions in its 3 April report, the review committee criticized the Hubble Space Telescope for a lack of transparency in reporting operating costs, and ordered Fermi, a γ -ray telescope, to cut costs by 10% each year from 2014. Only Spitzer, an infrared telescope, will be phased out earlier than its mission leaders wanted, in 2015. See go.nature.com/r6jgjo for more.

Carbon capture

The United Kingdom has relaunched a £1-billion (US\$1.6-billion) competition that will offer funding to companies that build commercial-scale facilities to capture and store carbon dioxide emissions from power



FRANÇOIS NOSTEN

Malaria drug resistance spreading

Malaria parasites that are resistant to the most effective current treatment — drugs based on artemisinin — are spreading in southeast Asia. Resistance was first confirmed in western Cambodia, close to Thailand, in 2008. But it is also emerging 800 kilometres westward, along the border of Thailand and Myanmar — where

villagers are shown being screened for malaria. Resistance will reach rates reported in western Cambodia within 2–6 years, says a 5 April study on the effectiveness of artemisinin treatments in more than 3,200 patients during 2001–10 (A. P. Phylo *et al.* *Lancet* <http://doi.org/hsu>; 2012). See go.nature.com/wefjya for more.

plants, to be in operation by 2016–20. The contest was first announced in 2007, but the bidders all pulled out. The relaunch, announced on 3 April, includes contracts guaranteeing the sale of the plants' electricity, which removes some financial risk. The competition also includes a further £125 million for research. See page 151 for more on carbon-capture plans around the world.

US public health

A committee convened by the US Institute of Medicine has recommended a suite of actions to remedy what it calls “dysfunction” in the funding, organization and capability

of the US public health-care system. Apart from urging a doubling of federal health-care spending, to be funded by taxing medical transactions, the 10 April report also endorsed more research into the effectiveness and value of public-health strategies. The 2010 Patient Protection and Affordable Care Act authorized a programme of research focused on similar issues — but that legislation is currently being challenged in the courts.

BUSINESS

Solar insolvencies

German solar-cell manufacturer Q-cells — which was the world's largest maker

of solar panels just five years ago — has filed for insolvency. The company, based in Bitterfeld-Wolfen, lost €846 million (US\$1.1 billion) last year and had been trying to restructure its debts, but gave up on 3 April. Three other German solar firms have filed for insolvency since December, as state subsidies decline and the market is flooded with low-cost photovoltaic modules and polysilicon raw material, much of it from China.

Imaging Alzheimer's

The US Food and Drug Administration on 6 April approved the first diagnostic test for plaques in the brain,

which are often associated with Alzheimer's disease. The test, developed by Eli Lilly, a pharmaceutical firm based in Indianapolis, Indiana, relies on a compound called florbetapir (Amyvid), which binds to amyloid plaques and can be imaged in living patients (see *Nature* **469**, 458; 2011). Florbetapir assays will complement existing behavioural diagnoses of dementia, but researchers hope that the tests will eventually become a tool for early diagnosis and treatment of the disease.

RESEARCH

GM pigs on hold

A genetically modified pig intended for human consumption seems unlikely to reach the dinner table any time soon, after local hog producers decided to pull their funding for the project. The Enviropig, developed at the University of Guelph in Canada, contains a transgene that enables it to better absorb phosphorus from its food, in turn reducing the phosphorus content of its manure and counteracting problems such as algal blooms in waterways fertilized by phosphorus run-off. It has not yet been approved for human consumption, and Ontario Pork, a Guelph-based group for hog producers that had

put around Can\$1.3 million (US\$1.3 million) into the project, said last week that it would not provide any further funding because the genetics had been proven; interest from industry would be required to restart the research.

PEOPLE

Anti-doping row

Complaining of restrictions to his freedom of speech, an anti-doping researcher has resigned from the Swiss panel that oversees biological passports — biochemical profiles of athletes that help to detect doping (see *Nature* **475**, 283–285; 2011). Michael Ashenden, who heads the Science and Industry Against Blood Doping Research Consortium in Gold Coast, Australia, said that he quit the Athlete Passport Management Unit in Lausanne because it added a confidentiality clause asking experts not to make public comments for eight years after they left the panel. The unit has just taken over the biological-passport system from the International Cycling Union in Aigle, Switzerland.

Heart-institute head

The US National Institutes of Health (NIH) has named cardiologist Gary Gibbons as director of the National Heart, Lung, and Blood Institute



(NHLBI). Gibbons (pictured) is currently at the Morehouse School of Medicine in Atlanta, Georgia, where he studies the genomics of vascular disease in minority populations. The NHLBI is the third-largest NIH institute, with a budget of roughly US\$3 billion, and has been without a permanent director since late 2009. Gibbons, replacing acting director Susan Shurin, plans to take up the post this summer.

FUNDING

Telescope rivalry

The 24.5-metre Giant Magellan Telescope (GMT), to be built in Chile, will not compete with its rival Thirty Meter Telescope in a contest for backing by the US National Science Foundation (NSF). With a mere US\$1.25 million for the winner, the contest was more about prestige than money. The NSF cannot

COMING UP

16–20 APRIL

The origin and evolution of life is the subject of a biennial gathering for astrobiologists, this year held in Atlanta, Georgia. go.nature.com/5blyic

16–21 APRIL

In Panama City, the newly formed Intergovernmental Platform on Biodiversity and Ecosystem Services lays out its priorities for monitoring global ecology. go.nature.com/o3zm54

provide significant cash for either ground-based telescope until at least 2020, although the winner might find it easier to attract international partners. Last week, the GMT's board of directors declared that the project could proceed on its own. See go.nature.com/rwc7df for more.

Swedish bioscience

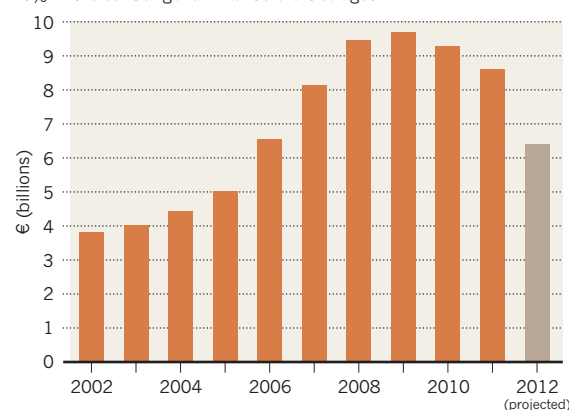
Sweden's government said on 3 April that the Science for Life Laboratory (SciLifeLab), an existing bioscience collaboration between four universities, will in 2013 expand into a national research institute for molecular biosciences and bioinformatics in Stockholm. It will eventually employ 1,000 scientists. The centre — which currently employs 300 scientists in Stockholm and Uppsala — is to receive 220 million Swedish kronor (US\$32.9 million) from the private Knut and Alice Wallenberg Foundation in Stockholm, and between \$25 million and \$50 million from pharmaceutical firm AstraZeneca. The government's contribution will be announced this autumn.

TREND WATCH

The Spanish government's austerity budget has lopped one-quarter off funding for research and development — much larger than the average 17% cut applied to other central-government departments. But public research funds may not be as badly hit as these figures suggest. More than half of overall funding relates to tax credits for technology transfer, for example, and last year almost half of those went unused. See go.nature.com/yalqgl for more.

SPAIN SLASHES RESEARCH SPENDING

Spending on research and development in Spain drops 26% in the central government's draft budget.



NEWS IN FOCUS

VIROLOGY Parsing the lessons of the mutant-bird-flu controversy p.150

BIOTECHNOLOGY Roche is in hot pursuit of medical sequencing p.152

PLANETARY SCIENCE A wet blanket on dreams of a watery Mars p.153



EPIDEMIOLOGY Chronicling a generation, in flesh, blood and data p.155

LIU HAIFENG/XINHUA PRESS/CORBIS



It's boom time for firms selling stem-cell treatments, such as those derived from umbilical blood.

REGENERATIVE MEDICINE

China's stem-cell rules go unheeded

Health ministry's attempt at regulation has had little effect.

BY DAVID CYRANOSKI

Three months after the Chinese health ministry ramped up its efforts to enforce a ban on the clinical use of unapproved stem-cell treatments, a *Nature* investigation reveals that businesses around the country are still charging patients thousands of dollars for these unproven therapies.

The clinics operate openly, with websites promoting the treatments for serious disorders such as Parkinson's disease, diabetes and autism, and attract thousands of medical tourists from overseas. They advertise case studies of individual patients who they say have benefited from the treatments, and some have clinics in major hospital complexes, giving them

an air of mainstream acceptance. Stem-cell experts contacted by *Nature* insist that such therapies are not ready for the clinic and say that some may even endanger patients' health. But the Chinese government is struggling to enforce its ban.

In May 2009, the Chinese Ministry of Health classified stem-cell treatments as Category 3 medical technologies, defined as "high risk" and requiring the approval of a technical audit board before use. So far, no approvals have been granted. Despite this, "one 2009 estimate put the number of stem-cell companies based in China at around 100", says Doug Sipp, a stem-cell ethics and regulation researcher at the RIKEN Center for Developmental Biology in Kobe, Japan. In his view, "even after the

reform efforts by the Ministry of Health, the industry apparently continues to grow."

In January, recognizing the worsening situation, the health ministry announced a package of rules for the industry. Organizations using stem cells must register their research and clinical activities, the source of the stem cells and ethical procedures. The ministry asked local health authorities to halt any unapproved clinical use of stem cells in their regions. And it called for a nationwide moratorium on new clinical trials for stem-cell therapies, adding that patients in existing clinical trials should not be charged.

So far, however, the ministry's clampdown has proved ineffective. According to a Ministry of Health spokesman, not one clinic has registered in the required way, and *Nature* has found that many stem-cell clinics continue to offer treatments. Shanghai WA Optimum Health Care, for example, which has plush headquarters in a gated estate in one of the wealthiest areas of central Shanghai, claims success in using stem cells derived from umbilical cord or adipose tissue to treat a range of disorders, from autism to multiple sclerosis. Tony Lu, a member of the company's science and technology board, says that four to eight injections of such cells can treat Alzheimer's disease, at a cost of 30,000–50,000 renminbi (US\$4,750–7,900) per injection. According to the company's senior patient-liaison officer, Karina Grishina, autism can be treated with an adipose-tissue-derived cell injection for 200,000 renminbi, followed a few days later by a 50,000-renminbi injection of umbilical cord cells.

In Changchun, Tong Yuan Stem Cell claims to have treated more than 10,000 patients with a variety of disorders, including Parkinson's disease. A representative says that it also offers a one-year, four-injection autism treatment protocol using stem cells from aborted fetuses. Meanwhile, Beijing Puhua International Hospital's Stem Cell Treatment Center offers a four-to-five-injection protocol for autism, costing 205,000 renminbi.

Those clinics all claim success in treating patients, but none has published data from controlled clinical trials. Zhou JingLi, chief neurologist at Beijing Puhua, says that many of the company's autistic patients have shown marked improvements in their condition a couple of weeks after treatment. She agrees that controlled clinical trials are needed to ►

► verify the efficacy of their treatments. “But,” she asks, “who’s going to pay?” She adds that first-hand experience with patients is enough to show that stem-cell treatment is worthwhile. “First, it’s safe; second, it’s effective. We know that,” she says.

Leading stem-cell scientists think otherwise. Commenting on Tong Yuan’s treatment for Parkinson’s disease, Oliver Cooper, director of the Stem Cell Facility of the Neuroregeneration Institute at McLean Hospital in Belmont, Massachusetts, and a specialist in Parkinson’s disease, says, “The products offered by Tong Yuan may provide anecdotal, poorly controlled, transient improvements in the patients, but Parkinson’s-disease patients need long-term therapies.”

“There are neither scientific nor clinical data to support the long-term benefits of haematopoietic- or neural-stem-cell therapies for Parkinson’s patients,” he adds. “In fact, it’s not clear if the infused cells will survive for more than a few days in the patients.”

Meanwhile, neurobiologist Ricardo Dolmetsch, an autism researcher at Stanford University in California, says, “The consensus in the autism research community, as well as in the stem-cell community, is that there is no scientifically valid reason for using stem cells

to treat autism spectrum disorders”. He worries that, without the proper safety studies in place, the treatments could “lead to serious complications like cancer and autoimmune disease”.

In addition to anecdotal case studies, some Chinese stem-cell companies bolster their reputations by claiming to have connections with leading politicians and scientists. A glossy ‘information memorandum’ from Shanghai WA Optimum Health Care contains pictures of staff with various local- and central-government figures, including Li Keqiang, the powerful executive vice-premier of the State Council who is tipped to succeed Wen Jiabao as China’s next premier.

It also lists Li Lingsong, director of the Peking University Stem Cell Research Center, as a member of its science and technology board. Li denies this. “I have so far nothing to do with WA,” he told *Nature*, adding that he has asked the company to remove his name. WA also claims a strategic alliance with Harvard Medical School, although neither the medical school nor the Harvard Stem Cell Institute is aware of any such connection. Likewise, the University of California, Irvine, where WA claims to have research facilities, denies any formal relationship.

When pressed, all of the stem-cell clinics approached by *Nature* said that they were aware of the government regulations, and that they were necessary — but only for other clinics that were not operating safely. Most emphasized that their own businesses were entirely legitimate. *Nature* did find one company, Shanghai Puhua, which says that it has already stopped offering stem-cell treatments to comply with government regulations. And Beijing-based Wu Stem Cells will probably do likewise for the same reason, says company director Cheng Bo.

Bioethicist Zhai Xiaomei at Peking Union Medical College in Beijing, a member of one of the government’s technical audit boards, was surprised to hear that any stem-cell companies were still operating. She says that the regulations are “absolutely clear” that companies must not administer unapproved stem-cell treatments.

A Ministry of Health representative told *Nature* that it was aware of the problem, and that it would be making greater efforts to clean up the stem-cell business.

Asked how WA operates despite the ministry’s ban, Lu describes the regulations as “a grey area”. Grishina agrees: “We’re in China, so there are different stipulations.” ■ **SEE EDITORIAL P.141**

BIOSAFETY

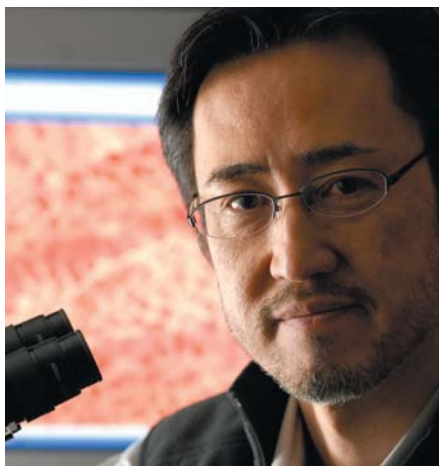
Post-mortem on mutant flu

Virus papers get green light but controversy highlights lack of global rules on biosafety.

BY DECLAN BUTLER

The dust is beginning to settle on the months-long controversy over two studies in which the H5N1 avian influenza virus was modified to be transmissible between mammals. But scientists and authorities still need to address the lack of international oversight for studies in which pathogens are deliberately made more dangerous, speakers emphasized at a two-day meeting held last week at the Royal Society in London. The meeting brought together scientists, research funders and experts in security, bioethics and foreign policy, just days after the US National Science Advisory Board for Biosecurity (NSABB) revised its earlier stance and recommended publication of the two studies.

In December 2011, the board recommended redaction of experimental details of the studies, on the basis of concerns about bioterrorism and the increased likelihood of accidental release of the viruses. But after considering revised versions of the manuscripts on 29–30 March, the NSABB voted unanimously



Yoshihiro Kawaoka can now publish his flu study.

in favour of full publication of the paper submitted to *Nature* by Yoshihiro Kawaoka of the University of Wisconsin–Madison, and 12–6 for publication of the content (although not the specific wording) of the paper submitted to *Science* by Ron Fouchier of the Erasmus

Medical Center in Rotterdam, the Netherlands.

Kawaoka presented his findings for the first time at the Royal Society meeting (see *Nature* <http://doi.org/hsn>; 2012), but Fouchier gave only a summary of his, saying that a more detailed description was prohibited by Dutch export-control laws, which require a permit to disseminate samples of, and information about, certain dangerous pathogens.

In December, the NSABB said that the information in the papers could help H5N1 surveillance efforts and so should be made available to experts on a need-to-know basis. A major factor in the board’s change of heart was that subsequent international discussions concluded that there was no practical way to selectively share the data, and that national export controls may restrict distribution anyway, says Michael Imperiale, a virologist at the University of Michigan in Ann Arbor and a member of the NSABB.

This left the board with the stark choice of publishing either all or none of the research, with publication becoming “the only way for public-health authorities to know what to look

J. MILLER/UNIV. WISCONSIN-MADISON

for", says Imperiale, who this time voted in favour of publishing both papers.

Some on the board were also influenced by clarifications to both papers — particularly those suggesting that Fouchier's viruses were not as pathogenic as they had initially seemed — and by presentations from Kawaoka and Fouchier indicating that some combinations of the mutations generated in their laboratories had already been seen in the wild. But Imperiale says that his vote was unaffected by the revisions. "I don't think the risks have changed; the authors have changed the host range and transmission properties of a deadly virus," he says. "I don't think the benefits have changed either."

Another factor in the NSABB's decision was the announcement on 29 March of a new US policy requiring that all publicly funded research on certain pathogens be assessed from the outset by the funding agency for the risk that it may be misused (see *Nature* <http://doi.org/hsp; 2012>). Experts have generally welcomed the guidelines, which they say should, if properly applied, also help to avert future repeats of the H5N1 controversy, in which the NSABB learned of the papers only shortly before their planned publication.

Arthur Caplan, a bioethicist at the University of Pennsylvania in Philadelphia, was one of several speakers at the Royal Society meeting to argue that it would be a "mistake" to think that the issues raised by the papers are now fully resolved. Besides bioterrorism, a major concern is that publishing them will result in worldwide proliferation of similar research, possibly in labs that may not have well developed biosafety cultures and training. Both Fouchier and Kawaoka worked in facilities rated at 'biosafety level 3 enhanced', but to expect that all such research would be done carefully everywhere is "utter malarkey", Caplan told the meeting.

The World Health Organization has recommended that work to make H5N1 viruses more transmissible in mammals, which flu researchers voluntarily halted in January, remain suspended until the relevant authorities have assessed the safety conditions for such research (see *Nature* **482**, 447–448; 2012). US and Dutch authorities are expected to release their verdicts within weeks, which Kawaoka said "would be the time to lift the moratorium". But others warned that doing so before broader debate, including planned hearings by US lawmakers, would be premature and could be perceived as arrogant.

The challenge for any oversight system will be to avoid discouraging important science while ensuring that work is limited to labs with appropriate safety standards, says one scientist, who criticizes the reluctance of many flu

researchers to admit that such studies carry risks. "You'd have to have your head in the sand not to accept that there are some risks here." ■

► **NATURE.COM**
For more see
Nature's mutant flu
special:
go.nature.com/mhmibi



Despite delays, China's GreenGen coal-fired gasification power plant in Tianjin is going ahead.

CLIMATE CHANGE

Slow progress to cleaner coal

China moves forward with demonstration power plant as United Kingdom revives carbon-capture programme.

BY JEFF TOLLEFSON AND
RICHARD VAN NOORDEN

With many of the world's nations dragging their feet on cleaning up fossil-fuel emissions, even slow progress stands out. This spring, China's state-owned Huaneng Group plans to fire up the first phase of its flagship clean-coal demonstration project, moving the country one step closer to capturing and storing the carbon it emits. Despite being more than a year behind schedule, the GreenGen coal gasification plant in Tianjin puts China at the forefront of global efforts to exploit coal resources without releasing carbon dioxide.

In 2008, leaders of the G8 group of nations called for the development of 20 large-scale projects demonstrating technologies for carbon capture and storage (CCS) by 2010, but countries have been slow to embrace the costly plants. Delays and cancellations have affected all but a handful of high-profile initiatives in Europe, the United States and Australia, whereas China, despite delays of its own, is still pushing forward to develop indigenous technologies.

"GreenGen represents both a high degree of technical sophistication and a real commitment on China's part to clean-energy

technology," says Julio Friedmann, head of the carbon-management programme at Lawrence Livermore National Laboratory in Livermore, California. "There can be no doubt that China has achieved something remarkable."

Originally estimated to cost US\$1.5 billion, GreenGen is being developed by a consortium of Chinese companies, including Huaneng, together with Peabody Energy of St Louis, Missouri. The first phase is a 250-megawatt integrated gasification combined-cycle power plant, which will convert coal into 'syngas' — a mixture of carbon monoxide and hydrogen — to be burned in specialized turbines to produce electricity. Waste carbon dioxide from these processes can be separated more easily than in conventional coal-fired power plants.

Huaneng has already begun work on a second phase — a smaller pilot plant that will send a clean stream of hydrogen through fuel cells and turbines to produce electricity, with carbon dioxide being captured for industrial use. The third phase, scheduled for 2015–20, will be a 400-megawatt power plant with full-scale carbon capture and storage in underground rock layers. That represents a substantial delay beyond the original completion date of 2015. Huaneng officials say that they revised the schedule in response to technical issues and delays to ►

► CCS projects in other nations.

GreenGen was originally seen as a follow-up to FutureGen, the flagship US initiative that was promoted and then cancelled by former president George W. Bush. In 2009, President Barack Obama revived the project, but a business consortium has yet to settle on a home for it. Many are sceptical about whether FutureGen and other US CCS projects will come to fruition, given the inability of Congress to craft a comprehensive programme for reducing greenhouse-gas emissions.

The United Kingdom is also trying to renew interest in CCS. On 3 April, the UK Department of Energy and Climate Change relaunched a £1-billion (US\$1.6-billion) competition for companies to build CCS demonstration plants. An earlier initiative, launched in 2007, fell apart in 2011 when the last remaining consortium pulled out because of concerns over costs. This time, the UK government has encouraged companies by promising that electricity produced by CCS plants can be sold at premium prices, and it will also provide £125 million for CCS research and development.

Many countries still see potential CCS plants as one-off demonstration projects, but the United Kingdom is trying to create conditions for them to become commercially viable, says Stuart Haszeldine, a geologist working on CCS at the University of Edinburgh, UK. So far, seven consortia have publicly announced their interest in the contest, which aims to commission plants in 2016–20 and store carbon dioxide in porous rock under the North Sea.

Other well-advanced initiatives include Canada's Boundary Dam Integrated Carbon Capture and Storage Demonstration Project in Saskatchewan, a 100-megawatt project to retrofit carbon-capture technology to an existing power plant; and the Texas Clean Energy Project, a 400-megawatt integrated gasification combined-cycle plant that could begin operating in 2014. Both plan to sell carbon dioxide to oil companies, which use the gas to flush oil out of reservoirs. In the Netherlands, the Rotterdam Capture and Storage Demonstration project aims to begin storing carbon dioxide in depleted gas fields under the North Sea in 2015.

Howard Herzog, a carbon-storage expert at the Massachusetts Institute of Technology in Cambridge, says that the challenge facing the entire carbon-capture industry is how to turn a profit in the absence of a serious carbon policy that rewards emissions-reduction projects financially. "The question is whether there

is a CCS market anywhere in the world," Herzog says. "If there isn't, why develop the technology?" ■

► **NATURE.COM**
For more about
GreenGen, see:
go.nature.com/2cmgm5

BUSINESS

Roche chases stake in medical sequencing

Biotech firm Illumina continues to resist takeover — but analysts suggest that a merger is inevitable.

BY ERIKA CHECK HAYDEN

Gene sequencing has mainly been the province of technology companies catering to researchers, but the pharmaceutical giant Roche, based in Basel, Switzerland, has other plans. It is increasing the pressure on Illumina, a DNA-sequencing company headquartered in San Diego, California, hoping to absorb the firm and use its expertise to capture the growing market in personalized medicine.

Roche proposed a merger in January, offering US\$44.50 per share, which Illumina's board rejected. On 29 March, Roche raised its bid to \$51 per share, for a total acquisition price of \$6.7 billion. Illumina's board rejected this offer, too, and adopted a 'poison-pill' provision that would give present investors the right to buy new shares at a reduced price, thereby diluting the value of Roche's shares if the deal goes through. But many observers say that the merger is all but certain. "If Roche really wants this, they are ultimately going to be able to make this go through," says John Haggerty, a mergers-and-acquisitions lawyer with Goodwin Procter in Boston, Massachusetts.

Roche's interest in Illumina stems from the latter's dominant position in genetic research; it claims that 90% of the world's sequencing is done on its machines. Roche believes that a whole-genome sequencing is poised to become a lucrative diagnostic tool, and that Illumina will allow it to break into that market.

Illumina could benefit from Roche's vast sales

force, which would make it more competitive with its major rival, Life Technologies of Carlsbad, California. And Roche's experience in diagnostics and drug development could speed the translation of sequencing into the clinic, says David Ferreiro, a Boston-based analyst at the investment bank Oppenheimer. Manufacturers are eager to move from next-generation sequencing technology into diagnostic technology, and "Roche could do that a lot faster" than Illumina could, he says.

Yet Illumina has fought back, arguing to its shareholders that Roche's offer undervalues the company (see 'Down and up'). Illumina, which started out in the 1990s as a microarray vendor, came to dominate the sequencing market by steadily cutting the cost and improving the speed and ease of use of its products.

However, Illumina now has numerous competitors with new platforms nipping at its heels. The firm is cagey about what technologies could replace its current sequencing machines, even as the newcomers, such as Oxford Nanopore Technologies in Oxford, UK, are promising big leaps in the speed and price of sequencing — with the aim of attaining the \$1,000 human genome. Industry observers question whether Illumina's nimbleness will survive a Roche merger. They point to Roche's 2007 acquisition of sequencing company 454 Life Sciences in Branford, Connecticut, which has since failed to live up to hopes that it would capture a significant part of the market.

To up the pressure, Roche has nominated six people for Illumina's board of directors who, if elected by shareholders at the company's annual meeting on 18 April, might coax the board to work out a deal. But with advisory firms cautioning major shareholders to hold out, Roche will probably continue to raise its price to entice institutional investors and arbitrage traders, who bought Illumina shares expecting Roche to sweeten its offer. "It's just a matter of how much the Illumina board can force them into raising their price before the stockholders start accepting the offer," says Haggerty.

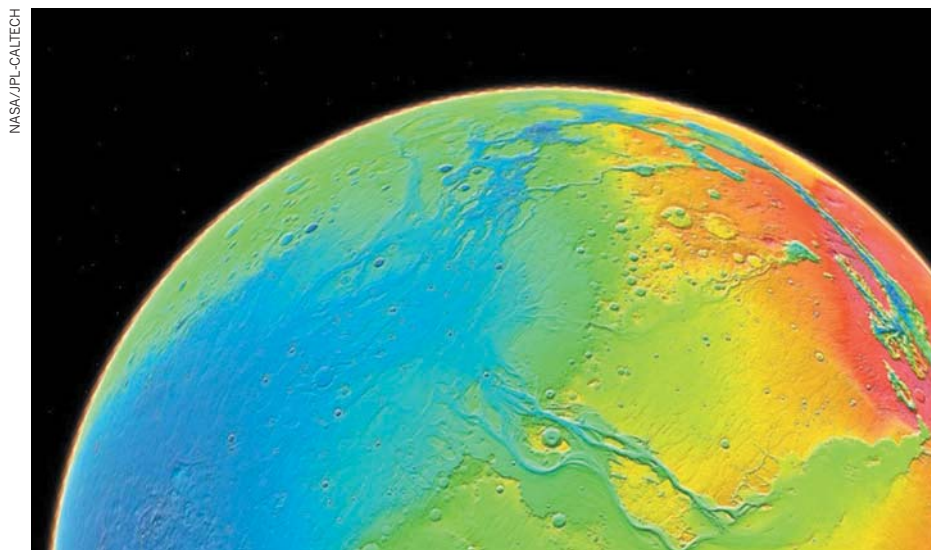
For Illumina, the end could be bittersweet. The company may be in a prime position to dominate the potentially lucrative clinical-sequencing market. But it may also see the end of its glory days as a key innovator in sequencing. ■

DOWN AND UP

Illumina's share price fell last year owing to gloom over biomedical-research funding. Roche's January bid boosted it, but only part way.



SOURCE: GOOGLE FINANCE



Sea change: previous ideas that lowlands on Mars (blue) once hosted oceans are being overturned.

ASTRONOMY

Dreams of water on Mars evaporate

Climate models reveal the red planet was mostly cold and dry.

BY ERIC HAND

The debate began when nineteenth-century Italian astronomer Giovanni Schiaparelli thought he saw water-filled *canali*, or channels, on the red planet: just how wet was Mars? “This is a pendulum that has been swinging back and forth,” says Jeff Andrews-Hanna, a planetary scientist at the Colorado School of Mines in Golden. The *canali* were an illusion, and no one doubts that Mars today is dry except for possible meagre seeps of groundwater. But in recent years researchers have come to accept that ancient Mars had lakes or even oceans — favourable conditions for life. “This is what started off the fever for Mars science,” Andrews-Hanna says.

But the pendulum is swinging again. Last month, Jim Head, a planetary scientist at Brown University in Providence, Rhode Island, threw a wet blanket on the idea that Mars was ever very wet at all, in a keynote talk at the Lunar and Planetary Science Conference in The Woodlands, Texas. Head and others are assembling a picture of a Mars that was cold and dry from the beginning, punctuated at most by short bursts of wetness. “The notion of a palm-tree-covered Mars has waned,” says Stephen Clifford, a planetary scientist at the Lunar and Planetary Institute in Houston, Texas, who is organizing a conference in May on the early climate of Mars.

The first spacecraft to visit Mars, in the 1960s and 1970s, showed a bone-dry planet pocked with craters, much like the Moon. But high-resolution cameras on orbiting missions such as the Mars Global Surveyor, which launched in 1997, showed valley networks — braided and branched channels about 3.7 billion years old that seemed to have been carved by water. Then in 2005, a spectrometer on the Mars Express satellite found widespread evidence for clays¹ — minerals that testify to hundreds or even thousands of years of exposure to water. Suddenly, geologists did not seem so bold in tracing out the palaeo-shorelines of an ocean that could have covered the planet’s entire low-lying northern hemisphere.

But Head and others have countered that view with three main lines of evidence. The first comes from models of the ancient Martian climate that fail to predict temperatures high enough for rain, or for liquid water to persist on the surface at all. The young Sun was fainter than it is today, and even if the young Mars had a thicker atmosphere, its greenhouse effect would probably not have warmed the planet above freezing, says François Forget at the University of Paris. He has submitted to the journal *Icarus* what he says is the most sophisticated three-dimensional

climate model for Mars so far. It predicts that any water on Mars would have been bound as ice at higher elevations.

By next month’s conference, Forget hopes to incorporate the effect of greenhouse gases, such as sulphur dioxide, that could have been present occasionally in huge doses. He hopes to test the idea, proposed by Head, that sulphurous bouts of volcanic activity could have warmed the atmosphere for brief periods, just enough to melt the icy highlands and unleash torrents that could have carved the valley networks. Other researchers invoke local melting from the heat from large asteroid impacts.

Closer inspection of the valley networks supports the sporadic presence of water rather than an enduring wet climate, Head says. He points to studies, including his own, showing that some of the networks are isolated geographically and in time, having formed hundreds of millions of years apart².

Even the clay minerals may not support a wet planet. A team using a spectrometer on the Mars Reconnaissance Orbiter found that roughly 80% of the clays occur together with other minerals that form at relatively high temperatures. This suggests that the clays formed not in cool surface water but underground, in water warmed by leftover heat from Mars’s formation, says Bethany Ehlmann, a planetary scientist at the California Institute of Technology in Pasadena, who led the study³.

Curiosity, the NASA rover that is expected to land on 5 August, may get a closer look. A thin ring of clay encircles a 5-kilometre-high mound in the middle of the Gale crater, where the rover will touch down. Although Ehlmann says that these clays are probably among the 20% that formed with surface water, their textures will reveal something about the extent of that water. Were the clays deposited in a persistent, deep lake? Or could the waters have been shallow and short-lived? Some geologists have even suggested that the minerals could have formed in the presence of ice.

Andrews-Hanna says the shift in thinking doesn’t rule out life on ancient Mars, but instead drives it deeper underground. “If Mars’s climate was never stable, that would have been a challenge for life,” he says. “But as you go deeper in the subsurface, things become more stable.”

Head doesn’t think his revisionism causes too many problems for ancient life even at the surface. He says ancient Mars could have been very much like the dry valleys and ephemeral lakes of Antarctica, where dried-up mats of algae bloom in the warmer season, nourished by a trickle of meltwater from the icy highlands above. “The public perception is that warm and wet equals life,” he says. “If you look at the range of life on Earth, there’s no reason to suspect that life is limited to that.” ■

NATURE.COM
For more on the red planet, see:
go.nature.com/tizth3

1. Poulet, F. *et al. Nature* **438**, 623–627 (2005).
2. Fassett, C. I. & Head, J. W. III *Icarus* **195**, 61–89 (2008).
3. Ehlmann, B. L. *et al. Nature* **479**, 53–60 (2011).



The Square Kilometre Array would be the world's most powerful radio telescope (artist's impression).

ASTRONOMY

Giant telescope may get two homes

Split-site solution could allow both Australia and South Africa to host parts of the Square Kilometre Array.

BY GEOFF BRUMFIEL

With the battle to host the world's most powerful radio telescope growing increasingly acrimonious, the project's leaders are considering whether to divide the spoils.

The US\$2.1-billion Square Kilometre Array (SKA) would open a window on the early Universe. As yet, international partners have not committed to covering the hefty price tag. But if the project goes ahead, it would bring a

flood of funding, prestige and scientific opportunities to one of the two competing teams: South Africa or joint bidders Australia and New Zealand.

Last month, after considering the scientific merits of the two sites, the SKA Site Advisory Committee concluded that South Africa offered marginally better opportunities (see *Nature* <http://doi.org/hst; 2012>). Since then, the already-intense lobbying from both sides has increased, with politicians from each country insisting that they would prevail.

Now, the SKA management board has asked a new scientific panel to determine whether the telescope, made up of 3,000 15-metre-wide dish antennas and many more simpler antennas, could be divided between the two proposed sites.

Politicians in Australia and South Africa say that they oppose any split in the project, but John Womersley, head of the SKA board and chief executive of the UK Science and Technology Facilities Council, says that a compromise may be one way to resolve the battle. "I have heard astronomers that I respect say that such a solution is possible," Womersley told *Nature*.

Astronomers plan to use the SKA to measure radio signals at a frequency of around 1 gigahertz, probing early galaxy formation and investigating how gravity behaves near black holes. Capable of detecting a television signal from up to 15 parsecs (50 light years) away, the telescope might even aid in the search for extraterrestrial intelligence. Supercomputers will integrate the signals from the dishes and antennas, creating a virtual dish with a combined collecting area of 1 square kilometre (see *Nature* **480**, 308–309; 2011).

"Normally when you have a giant dish, you cannot split it, but the SKA has many different components," says Heino Falcke, a radio astronomer at Radboud University in Nijmegen, the Netherlands. The easiest solution would be to divide the project by placing the higher-frequency dishes on one continent and the lower-frequency antennas on the other.

Doing so would almost certainly raise the SKA's price tag, because computing centres and power would be needed in both remote locations, says Albert Zijlstra, director of the Jodrell Bank Centre for Astrophysics in Manchester, UK. But splitting the antennas by frequency would avoid the need for a high-throughput data link connecting the two sites, something that Zijlstra and Falcke both expect would be even more costly.

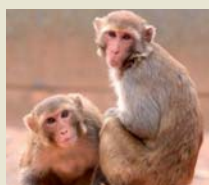
Zijlstra says that he can see little scientific advantage to splitting the project, aside from a few extra hours of observation time gained from having the telescope's two parts separated by such a vast distance. Falcke agrees. "It's a question of politics," he says.

The panel is expected to deliver its results by mid-May, when the SKA board will meet again to discuss the site. ■

SWINBURNE ASTRONOMY PRODUCTIONS/SKA PROGRAM DEVELOPMENT OFFICE

 **MORE ONLINE**

TOP STORY



Social status changes gene expression in monkeys
go.nature.com/mkjjbo

MORE NEWS

- Magnetic storms spotted on Venus
go.nature.com/l4wqgw
- Astronomers find solar system more crowded than ours
go.nature.com/s4ejqr
- United Nations to appoint a chief scientific adviser
go.nature.com/tnk229

Q&A



Private-spaceflight pioneer Elon Musk sets his sights on Mars
go.nature.com/w99pls

COMING OF AGE

Researchers in Britain have tracked thousands of children since their birth in the 1990s. Now the study is 21, and turning to the next generation.

BY HELEN PEARSON

In a secure storage barn on the outskirts of Bristol, nearly 9,000 placentas float in plastic buckets of formaldehyde.

Twenty-five kilometres away, in the basement of a university building, baby teeth from more than 4,000 children fill cardboard boxes in a walk-in freezer. Next door are some 15,000 nail clippings and 20,000 locks of hair. A few steps farther, a parade of freezers house row upon row of bar-coded blood cells, plasma, urine, saliva and chunks of umbilical cord that together make up a tissue library with more than one million entries.

This library is the harvest from an unusual study of humanity. In 1990, researchers started to collect tissues and detailed information from more than 14,500 pregnant women in this western British city and its surrounding region of Avon. The women filled in more than 100 pages of questionnaires about their health, relationships, work and home. After birth, researchers tracked the children's development through surveys, clinical examinations and biological samples. They know what the kids ate, when they first talked, how often they fell sick and when a parent read to them — or deserted them. They know when the children started to hit puberty, drink alcohol, have sex and leave home. In that wealth of data — collected at a cost of some £42 million (US\$67 million) so far — they are tracing how genetic and environmental factors in the children's early years affect their later ones.



Amy Murdoch-Davis is one of 14,000 people who have been studied since birth in and around Bristol, UK; her baby, Esmé, will soon join the study, too.

Now, the Bristol cohort study is coming of age, literally and scientifically. This spring, the first members of the group turn 21. And on 18 April, leaders of the study and their collaborators from around the world will meet in Bristol to discuss what they have learned from the Avon Longitudinal Study of Parents and Children (ALSPAC), also known as the Children of the 90s. The rich collection of tissues and behavioural data — a comprehensive phenotype for each of thousands of participants — makes the study unique and invaluable, say its leaders. “It’s the deepest phenotyping and biobank resource of any large birth cohort — unequivocally,” says George Davey Smith, the study’s scientific director. The results have generated more than 700 scientific papers and range from policy-changing health advice for pregnant women and young children to the discovery of genetic factors involved in fetal growth, obesity, allergies and bone density^{1–4}. The study has also inspired and guided other birth cohorts, including the world’s largest,

which is tracking more than 100,000 children in Norway.

Yet many of the data and samples remain untouched, and the cohort leaders acknowledge that the most important findings are likely to emerge years from now, some of them courtesy of new techniques that the founders of the study could barely have imagined. Davey Smith and his colleagues are just starting to analyse the children’s genomes — 2,000 of which have been fully sequenced — and epigenomes, the catalogue of chemical footprints left on a child’s DNA by experiences in the womb or in the world. The researchers say that such studies will help them to move from a slew of loose epidemiological associations — between a mother’s fish-eating and her child’s intelligence, for example — to the genetic and epigenetic players responsible for those links.

As the children themselves become parents, the team is expanding the scope of the study to a new generation. Still, at the heart of all this remains an unanswered question: how much value is there in collecting a huge amount of

ALSPAC (CHILDREN OF THE 90S)



The Bristol study has banked more than 1 million samples including blood, urine, saliva and placentas.

information about human life when the scientific pay-off is unknown? Marcus Pembrey, who was director of genetics at ALSPAC until 2005, recalls that when he initially told colleagues that he wanted to measure “everything” about the children, “they laughed and said you can’t study everything. And I said you have to study everything. You don’t know what will be important in the future.”

CONCEPTION AND BIRTH

Today, the effort to gather everything is headquartered in a concrete slab of a building in central Bristol. On the ground floor, a brightly painted annex houses a clinic where the children of the 90s come for their regular medical examinations. A few middle-aged men — the dads of the cohort members — are here, going from room to room for blood-taking, cognitive testing and bone scans. The study scientists are bringing in as many dads as they can to collect clinical data such as DNA, height, weight and blood pressure, in an effort to ramp up studies on the fathers’ health, as they already have for the mothers.

Also in the clinic is 19-year-old cohort member Amy Murdoch-Davis, with her five-month-old baby Esmé. Last month, ALSPAC received provisional ethics approval to start recruiting all the children of its cohort members, an effort it hopes to launch in earnest later this year. Esmé will be one of the first signed up. “I’ve been a guinea pig all my life,” Murdoch-Davis says, “and she can be a guinea pig too.” This time around, the researchers want to collect even more samples, including breast milk and the babies’ first stools.

The study itself was a struggling infant when Murdoch-Davis was born. Its leader, Jean

Golding, a mathematician turned epidemiologist, had studied stillbirths and neonatal deaths and was eager to learn how events in pregnancy and infancy affect a child’s health and development. She was convinced that studying a large cohort, starting in pregnancy and gathering as much information as possible, was the best way to find answers. Funders, however, saw it differently. “Everybody thought ‘if you have a cohort study you’re tied down to carrying on funding it — and it’s a bottomless pit,’” she says.

Golding eventually scraped together some money by writing grant applications that focused on specific diseases. She drummed up interest from mothers by talking on television and radio, and sending an army of midwives to antenatal classes and doctor’s surgeries. The 14,541 pregnancies eventually included in the study encompassed more than 70% of those eligible in the region between April 1991 and December 1992. The buckets of fresh placentas started stacking up. And so did other data, such as blood squeezed from the babies’ heels during the early clinic visits and reams of questionnaires filled out by the mothers (see ‘Building a bank of life’). Little was left unasked. Does baby drink breast milk or formula? Has he or she had antibiotics or skin ointment? Do you have a telephone; a tumble dryer; cockroaches?

Within a year or two of starting, however, the cash was draining away. Golding employed her 40–50 staff one month at a time, never

sure whether she would have enough to pay them again. And there was no time or money to analyse the data that were flooding in. “We were well in the red,” Golding says. “I was just exhausted.”

CHILDHOOD AND ADOLESCENCE

A few years later, just as parents were collecting baby teeth from beneath thousands of pillows and mailing them to the research team, the first data analyses started coming through. Researchers now point to a handful of results for which the cohort is famous, and that had an impact on public-health policy. One showed that eating oily fish during pregnancy was associated with better eye and cognitive development in children^{5,6}. Another helped to cement advice that babies should be put to sleep on their backs to reduce the risk of cot death, by showing that this sleeping position did not cause any developmental delays⁷. A third showed the first association between peanut allergy — an emerging epidemic in Western countries — and peanut oil in baby lotions⁸. Manufacturers soon started identifying the ingredient on labels.

As the children entered their second decade, the study found itself on firm financial ground for the first time. The Wellcome Trust and the Medical Research Council awarded it core funding that has totalled some £21 million for 2001–14, and investigator-led grants have brought in much more.

With the study maturing comfortably, Golding retired in 2005. (“They stopped paying me,” as she puts it, still firmly ensconced in her office at the University of Bristol.) Lynn Molloy, a social scientist, took over the managerial reins, and Davey Smith, an epidemiologist passionate about genetics, grasped the scientific ones.

The study entered a productive scientific adolescence even as the cohort members suffered their own teen agonies. Questionnaire

data revealed that 19% of 16–17-year-olds cut or otherwise hurt themselves. By age 18, some 5–10% had experienced some form of psychosis — “more common than people had previously suspected”, says Stanley Zammit, a psychiatrist working with cohort data at Britain’s Cardiff University — even though few are likely to go on to develop schizophrenia or a related condition.

At the same time, genetics was finally entering the picture. In the early years of the project, examining the children’s genotypes was just a fond hope. Scientists had announced a draft human genome in 2001, but that was a multibillion-dollar megaproject. Quite what to do with several thousand raw human genomes,

“IT’S THE DEEPEST PHENOTYPING AND BIOBANK RESOURCE OF ANY LARGE BIRTH COHORT.”

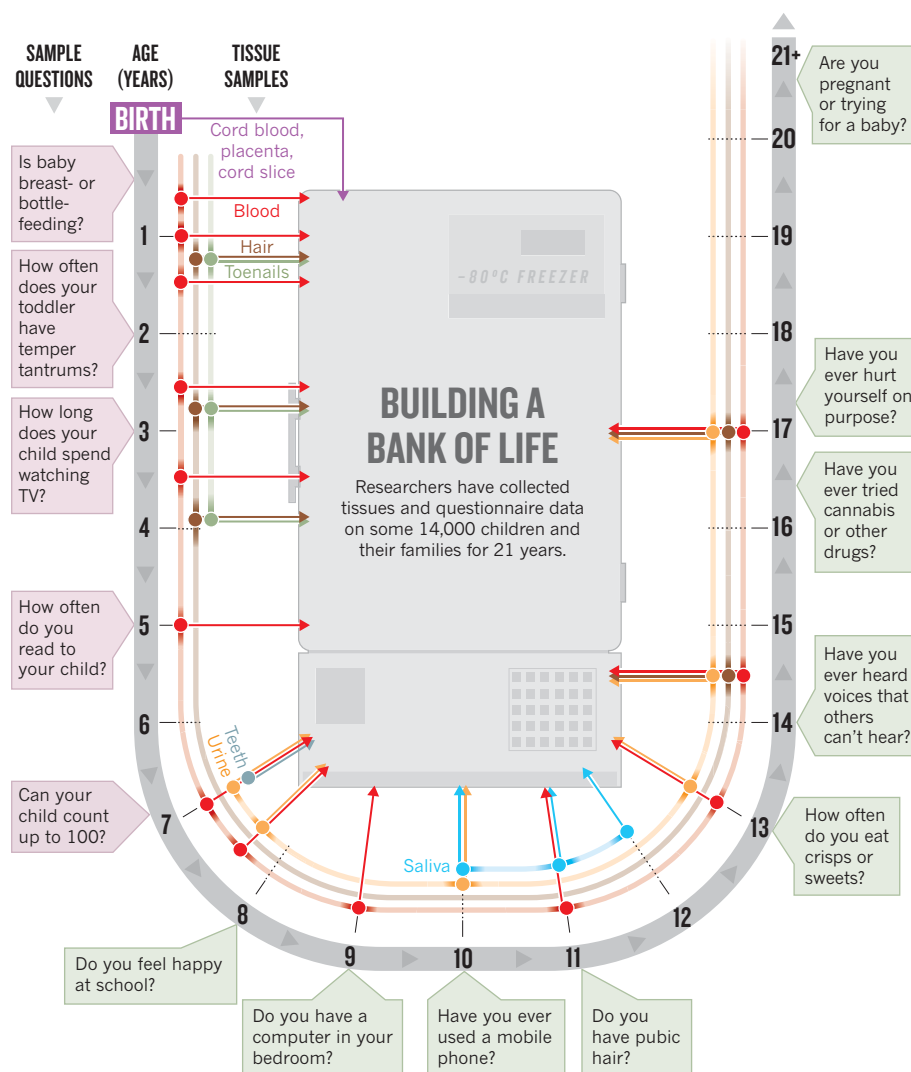
no one knew. “The interesting thing was that although all these high powered alpha-male geneticists were talking the big talk about genetics — when you actually gave them DNA for a thousand individuals it was too much. They didn’t have the technology,” Golding says.

In 2007, however, data from ALSPAC and several other large human biobanks were used to scour the human genome for single-letter variants associated with obesity. The work turned up a gene called *FTO*, and found that adults with two ‘risk’ copies of this gene are about 3 kilograms heavier, on average, than those with no risk copies². The discovery became a poster child for the identification of risk alleles through genome-wide association studies, a technique that was sweeping through the genetics community at the time. The Bristol cohort — with its extensive bank of DNA and medical data — was perfectly placed to catch the wave. Its data have been used to identify genetic sequences associated with fetal growth¹, bone density and childhood growth⁴, tooth development⁹, facial features¹⁰ and more. Researchers are now hunting for genetic links to intelligence, educational attainment and gender orientation.

Davey Smith sees other opportunities to explore associations in the cohort’s DNA. He finds it frustrating that epidemiological studies often reveal a correlation between two factors — poverty and obesity, say — without proving that one causes another. The trouble is, many social and biological factors tend to correlate anyway: people who smoke also tend to drink more alcohol, eat unhealthy food, be poorer, weigh more and have high cholesterol and other signature biomarkers of ill health.

One way to filter true causes from the correlations is to compare one cohort with another, something Davey Smith did recently to find ‘causal associations’ with breastfeeding. In the Bristol cohort, breastfeeding is correlated with less obesity, lower blood pressure, higher intelligence and more good things besides. But in the United Kingdom, breastfeeding mothers are also more likely to be middle- or upper-class. So is it breastfeeding that helps children, or some other aspect of a comfortable life? When Davey Smith and his colleagues looked at a birth cohort based in Pelotas, Brazil, where breastfeeding is not linked to socioeconomic status, the links with obesity and blood pressure fell apart, but the link with IQ held up¹¹.

Researchers are also finding true causes through Mendelian randomization, a genetic technique that is exciting epidemiologists. Stephanie von Hinke Kessler Scholder, an economist at the University of York, UK, is using this method to unpick the associations between children’s obesity and poorer performance in school. Obesity is also correlated with lower socioeconomic status, so Scholder sought to test whether obesity directly hinders performance (because of bullying, for example), whether kids who are obese do less



well because they come from disadvantaged families, or whether some other factor could explain the correlation.

Scholder took the banked DNA of the ALSPAC children and divided them up into groups on the basis of the make-up of two weight-related genes, *FTO* and *MC4R*. (Children with all ‘heavy’ copies of these genes weigh a few kilograms more on average than those with all ‘light’ copies.) Gregor Mendel’s laws of inheritance ensure that the heavy and light forms of the genes are shuffled and randomly delivered to children across a population, irrespective of their social class or any other confounding factor. And when Scholder removed confounders from the equation in this way, she found that children with the heavy versions of *FTO* and *MC4R* did just as well on school tests at age 14 as those who didn’t have them. Dispelling the false idea that fat children do worse at school is “a positive thing”, she says¹².

The most technologically advanced answer to the causation problem sits upstairs in the Bristol headquarters of the study. There, a £300,000 machine is poised to start rapid analyses of methylation — an epigenetic mark that

controls gene activity. By looking at 450,000 sites in the cohort members’ genomes, the researchers will build up a bank of methylation data on blood samples taken from the children at birth, at ages 7 and around 17, and on the mothers during pregnancy and 17 years later. “It’s a unique resource,” says Davey Smith.

In a study published last month¹³, Caroline Relton at Newcastle University, UK, who is leading the epigenetic work for ALSPAC, looked at methylation on an array of genes in the umbilical-cord blood of cohort children. She found that methylation signatures on nine genes at birth showed some association with body height and tissue composition at age nine. The finding hints that events during pregnancy might shape gene expression early in life — eventually resulting in altered growth and weight gain. It also “provides a flavour of the sort of thing we can address more powerfully when we get this huge data set under way”, Relton says. Next, the team plans to search for methylation patterns linked to factors such as neural development, behavioural problems, cardiovascular health and fatty liver disease. The researchers will



Jean Golding, pictured here with cohort children in 2001, started the study in the early 1990s.

also look for potential causes of these methylation patterns — associations with a mother's smoking, alcohol intake, weight gain during pregnancy, folate levels and exposure to air pollution.

NEXT GENERATION

The methylation robot is just one piece of the high-tech equipment involved in the study. In one lab in Bristol, blood cells taken from the dads that day are being spun down and divided up into aliquots, ready for storage. In another, researchers are transforming cell samples into immortal cell lines. They have already banked around 7,000 such lines, which provide an endless supply of DNA and might be used for future studies of cell behaviour. Sue Ring, the head of the laboratories, says that her team take turns to carry an emergency mobile phone that will warn them of a freezer failure. After all that the participants have given to the study, she says, she feels a "duty of care for the samples".

Soon, the labs will start to process a whole new set of tissues — from the children of the cohort members. The eggs that will develop into these children were formed when their mothers themselves were babies, growing in their own mother's uterus. This means that events in the grandmothers' lives, such as an infection, stress or exposure to toxic chemicals — all recorded in the cohort database — could produce signals in the streams of data that the researchers plan to collect about

the grandchildren. "I was going to retire and wind down until we stumbled on these trans-generational events," says Pembrey, who wants to explore the third-generation consequences of traumatic events in the lives of grandparents.

With so much yet to do, has the study justified the years of effort put in so far? Cohort studies sometimes draw criticism because "they are very big and very long-term and they seem to be a lifelong source of income for investigators", says Teri Manolio, who has worked on cohorts in the past and is now director of the Office of

Population Genomics at the National Human Genome Research Institute at Bethesda, Maryland.

But ALSPAC receives kudos from many researchers, including Nigel Paneth, an epidemiologist at Michigan State University in East Lansing, who runs a data-collection site for the US National Children's Study, which started recruiting in 2009.

Unlike ALSPAC, the National Children's Study has ballooned into a mammoth programme with costs estimated at as much as \$6 billion, and has struggled for more than a decade to convince sceptical scientists and funders that it will pay off scientifically. Paneth says that big cohort studies can make major contributions to health if they are well planned and executed. "Can I guarantee results? Of course I can't. But I can guarantee you no results if we don't do it."

In Bristol, the team is only starting to learn the value of some of the data collected decades ago. The placentas, which had to be moved

twice, were regarded mostly as a "bloody nuisance" until three years ago, when David Barker at the University of Southampton, UK, and Oregon Health and Science University in Portland, asked to use them. Barker is famous for drawing connections between early fetal development and adult health; in the past few years he has reported that the size and shape of a placenta — the life support of the fetus — is associated with risk of adult coronary heart disease, high blood pressure and even lifespan¹⁴. He sent a photographer to Bristol to snap the placentas, and is now using the organs' dimensions to test some of his hypotheses. The correlations are "definitely there and amazingly strong", Barker says.

Other correlations will become clear only when the study and its participants have grown up a little more. They do, after all, have most of their lives ahead of them. Murdoch-Davis is planning the next step in hers: going to university to study English and psychology. Meanwhile, Molloy is mapping out what data to collect when the cohort members next visit the clinic, aged 24–25, when they will be close to the peak of their health. And Davey Smith has faith in future generations of scientists to find new ways to study the cohort. "My hope is that when I drop dead it's got an energetic scientific director who's implementing all these future technologies that I couldn't imagine," he says.

Perhaps, Davey Smith says, someone will be analysing data from the digital video recorders that he's thinking about handing to each expectant family, to document every significant event in their child's life. After all, storing terabytes of data doesn't cost anything — and you never know what they might one day reveal. "Digital recording allows you to store up huge banks of data for future use," he says. "It's a bit like Jean collecting the placentas. No one knew what they were going to do with them." ■

Helen Pearson is Nature's chief features editor.

1. Freathy, R. M. *et al.* *Nature Genet.* **42**, 430–435 (2010).
2. Frayling, T. M. *et al.* *Science* **316**, 889–894 (2007).
3. Henderson, J. *et al.* *J. Allergy Clin. Immunol.* **121**, 872–877 (2008).
4. Timpson, N. J. *et al.* *Hum. Mol. Genet.* **18**, 1510–1517 (2009).
5. Williams, C. *et al.* *Am. J. Clin. Nutr.* **73**, 316–322 (2001).
6. Daniels, J. L. *et al.* *Epidemiology* **15**, 394–402 (2004).
7. Hunt, L., Fleming, P. & Golding, J. *Pediatrics* **100**, E11 (1997).
8. Lack, G., Fox, D., Northstock, K. & Golding, J. *N. Engl. J. Med.* **348**, 977–985 (2003).
9. Pillas, D. *et al.* *PLoS Genet.* **6**, e1000856 (2010).
10. Paternoster, L. *et al.* *Am. J. Hum. Genet.* **90**, 478–485 (2012).
11. Brion, M.-J. A. *et al.* *Int. J. Epidemiol.* **40**, 670–680 (2011).
12. von Hinke Kessler Scholder, S. *et al.* Preprint available at www.bristol.ac.uk/cmpo/publications/papers/2011/wp274.pdf
13. Relton, C. L. *et al.* *PLoS ONE* **7**, e31821 (2012).
14. Barker, D. J. *et al.* In *The Placenta and Human Developmental Programming* (eds Burton, G. J. *et al.*) Ch 2, 5–12 (Cambridge Univ. Press, 2010).

COMMENT

ATMOSPHERIC SCIENCE Regulate China's air quality for global impact **p.161**

FICTION Michael Frayn on expressing uncertainty through farce **p.163**



HISTORY Thomas Kuhn's paradigm shift **p.164**

OBITUARY Sherry Rowland's battle to save the ozone layer from CFCs **p.168**

ILLUSTRATION BY JONATHAN BURTON



Bold strategies for Indian science

For a nation of its talent and education, India deserves higher scientific standing. It needs clear and honest leadership, not more money, says **Gautam R. Desiraju**.

When an Indian prime minister publicly admits that India has fallen behind China, it is news. Manmohan Singh's statement last January at the Indian Science Congress in Bhubaneswar that this is so with respect to scientific research, and that "India's relative position in the world of science has been declining", has rung alarm bells. Singh was not springing anything new on Indian scientists; many of us will admit that things are not well¹. Recognizing the problem is the first step towards reversing this slide.

At present, India has a trickle-down strategy, in which elite institutions are

supported in the hope that good science there will energize the masses, and a bottom-up approach, in which the general public is targeted with schemes to popularize science.

These approaches have converged with the setting up in recent years of tens of new universities, institutes and centres of higher learning, even though many hundreds more are desirable for a country of India's size. Although there was, curiously, no increased allocation to science in this year's Indian budget, there is hope that, as the prime minister has declared, things would improve if government support were increased to 2% of the gross domestic product (it now stands

at 0.9%). But it is a haphazard plan, with no hint of new strategies. The assumption is that the answer to our problems lies simply in more money.

As someone who has worked in India for 34 years, I am impatient with our slow progress². At the glitzy level, we have had no Nobel prize winner since C. V. Raman in 1930, no highly Shanghai-ranked university, no miracle drug for a tropical disease and no sequencing of the rice genome. At the industrial level, there have been no breakthroughs to rival the telephone, the transistor or Teflon. At the organizational level, we do not have a postdoctoral ▶

► system worth its name, and our undergraduate teaching system is in a shambles. We figure occasionally in the best journals, yet we tolerate plagiarism, misconduct and nepotism. And yet, the innate abilities and talents of India are palpable. Why is it that this country has not been able to harness its strengths into deliverables?

Money is not the primary constraining factor in our problems, nor will an abundance of it solve them. More money is undoubtedly better, but if there are deep cultural and social problems, extra money will simply drain away. The rate of any improvement will not match the rate of increase in investment³. Big problems in big countries usually emanate from a small number of core reasons. An understanding of these reasons in the context of Indian science should stem from an appreciation of the country's historical, economic and sociological profile.

It is not enough for the prime minister to resort to platitudes by saying (as in his recent speech) that “things are changing but we cannot be satisfied with what has been achieved”, or that we should make “scientific output more relevant”. He and his advisers must ask themselves if there are underlying causes for this lack of satisfaction and relevance. Until then, no amount of bankrolling, populism, bureaucrat bashing or whistle-stop tours by prominent Western scientists will help.

A FEUDAL MINDSET

Two aspects of the Indian psyche are particularly troubling for a country seeking its rightful place in the modern world. Our cultural value system, backed by Hindu scriptural authority, has created a strongly feudal mindset among Indians. Centuries of servitude, right up until 1947, have made the average Indian docile, obedient and sycophantic. ‘Behave yourself and be rewarded’ is the pragmatic mantra. I believe this feudal-colonial mentality has had far-reaching and debilitating consequences for research.

The first is our lack of the ability to question and dissent that is so essential to science. Most of the faculty in our better institutions have done postdoctoral work in a foreign laboratory of consequence. Unlike young scientists in advanced countries, however, newly returned Indian lecturers typically relive their golden moments as postdocs throughout their research careers. The best research papers from India may be competent, but they do not inspire or excite. Very few Indian scientists are known as opinion-makers, trend-setters or leaders. They follow obediently.

Another consequence of this feudal mindset is our unquestioning acceptance — bordering on subservience — to older people. In this part of the world, age is blindly equated with wisdom, and

youth with immaturity. This facilitates the continuance of the status quo. Geriatric individuals with administrative and political clout reinforce their positions so well that we are unable to eject them. So we hail scientists in their eighties, film actors in their seventies and cricketers in their forties.

VARIANTS OF CORRUPTION

In healthy organizational hierarchies, the decision-makers are also active participants who have a stake in the future. We will have come of age only when Indian universities are allowed to appoint their own vice-chancellors, and institutes and national laboratories their own directors, rather than suffer the choices made by conclaves of old men in New Delhi.

The most important decisions in an academic system concern the appointment of faculty. This procedure is flawed in India. For a start, selection committees consist mostly of outsiders, and representation from within is often restricted to institutional and departmental heads. In the smaller state universities, all sorts of irregularities occur in the name of caste-based reservations. In the more influential central institutions, appointments are often made incestuously, with students of a few senior researchers filling a disproportionate

“More money is better, but if there are deep cultural and social problems, it will simply drain away.”

number of vacancies, or with plain academic ‘inbreeding’. A good dose of regional parochialism completes the picture.

Corruption need not take a monetary form; in a national laboratory it can mean acquiescing to the notion that one's administrative head is also one's scientific superior. By that logic, and given our civilian-based system of administration, the secretaries in the science ministries in Delhi should be our most creative scientists.

These variants of corruption — along with general indifference, absence of incisive introspection, old-boys' networks, administrative vindictiveness, vagaries in research funding and studied silences — conspire to create an atmosphere that lacks innovation and creativity. Impact factors and *h*-indices become the sole arbiters of scientific excellence in such an environment. If policy-makers are ignoring cultural parameters, scientists are looking only at numerical parameters.

The true measures of a country's scientific strength are found in the numbers of competent teachers and lively students in schools and undergraduate colleges, because these translate into real gains in the future. Fluffy factors, such as the numbers of

articles in *Nature* and *Science*, do not tell the real story. As a chemist, I would say that it is better to move deliberately and confidently towards the thermodynamic minimum (a system's most stable state, which has the lowest energy but is not always the easiest to achieve) rather than flit anxiously between any number of metastable kinetic states (which are easy and fast to access but have higher energies).

A large country with a well developing economy can afford this long-term strategy and vision. China need not be a comparison point⁴ — India is endowed enough to seek its own solutions for its problems⁵.

THE WAY FORWARD

I suggest that our policy-makers consider the following. First, provide modest funding to a very large number of small, single-investigator, blue-sky projects — including those in state universities — to achieve a critical density of ideas and a feeling of mass participation and enthusiasm.

Second, provide heavy and directed funding into a few specific projects of national importance — such as energy, water and public health — with high levels of accountability and proper exit options. Third, reduce or abolish the present system of awards, prizes and recognitions in higher-level science. This would dissuade younger scientists from chasing awards rather than doing good science, and it would reduce the influence of the cliques who allocate prizes.

To reach a stable solution, we can employ longer-term measures that include modification or removal of caste-based quotas and reservations in the educational and research sectors; improvement of undergraduate teaching institutions and teaching laboratories with respect to greater uniformity and transparency; and clear identification of paths towards scientific and administrative growth for individuals.

Money is neither the cause nor the solution to our problems, although it can facilitate progress in an otherwise healthy climate. What is lacking in India is the quality of leadership and the level of honesty that are required for a breakthrough. When will this country see another C. V. Raman? ■

Gautam R. Desiraju is a professor of chemistry in the Indian Institute of Science, Bangalore 560 012, India. He is president of the International Union of Crystallography. e-mail: desiraju@sscu.iisc.ernet.in

1. Jayaraman, K. S. *Nature* <http://nature.com/doi/10.1038/nature.2012.9750> (2012).
2. Desiraju, G. R. *Econ. Polit. Weekly* **43**, 37–43 (2008).
3. Shi, Y. & Rao, Y. *Science* **329**, 1128 (2010).
4. Kostoff, R. N., Bhattacharya, S. & Pecht, M. *Technol. Forecast. Soc.* **74**, 1519–1538 (2007).
5. Desiraju, G. R. *Angew. Chem.* **123**, 5704–5705 (2011).



Rapid economic growth means that the air in some Chinese cities, such as Beijing, contains more fine particles than the World Health Organization recommends.

Cleaning China's air

To reduce airborne soot, organics and sulphates, tailored strategies for each must be established and coal use limited, say **Qiang Zhang, Kebin He and Hong Huo**.

On 29 February this year, China's State Council approved its first national environmental standard for limiting the amount of fine particles in air that measure less than 2.5 micrometres in diameter. It requires the country to implement the World Health Organization's recommended interim target of an annual average of 35 micrograms per cubic metre ($\mu\text{g m}^{-3}$) for such particles by the end of 2015.

Fine particles — including soot, organics and sulphates — have a severe effect on human health and are implicated in climate change. They are emitted by combustion and industrial processes, and formed from the reactions of gaseous pollutants. If implemented properly, China's air-quality standard would have far-reaching benefits: as well as protecting human health, it would reduce air and mercury pollution in the Northern Hemisphere and slow global warming.

Achieving this goal will be a challenge. Some Chinese cities currently have fine-particle concentrations that are well above the proposed standard: levels of more than $100 \mu\text{g m}^{-3}$ have been reported¹. To meet the ambitious air-quality limits, China will have to overcome two major hurdles: its relentless increase in fossil-fuel use, which quickly wipes out any efforts to reduce emissions, and its decentralized system of environmental enforcement, which gives

undue influence to local officials who favour economic development.

Controlling air quality in China will address global environmental issues. For example, pollutants from east Asia that travel across the Pacific increase ozone concentration in the western United States². This could be relieved if China reduces emissions of nitrogen oxides (NO_x), which are precursors of fine particles and ozone. Cross-border pollution by airborne particles would similarly be reduced by cutting China's emissions of sulphur dioxide (SO_2). The use of technologies such as desulphurization, selective catalytic reduction or electrostatic precipitators to reduce Chinese emissions of SO_2 , NO_x and fine particles, respectively, can remove global pollutants such as mercury, which is released by the burning of coal.

Limiting particle pollution will also affect drivers of climate change — but not always for the better. On the one hand, reducing soot emissions by cutting coal use or using cleaner stoves will lessen radiative forcing and thus limit warming, benefiting both the climate and public health³. A stricter emissions standard for diesel vehicles, which produce soot, is another win-win solution⁴. On the other hand, reductions in SO_2 emissions from power plants would reduce atmospheric sulphate concentrations, thereby increasing radiative forcing, which has a short-term

detrimental effect on climate³. Thought is therefore needed as to how the various pollutants and sources should be best controlled, and a multi-pollutant abatement strategy must be developed.

A CLEAN CHALLENGE

The control of air pollution in China is in a race with the economy. The country has maintained an annual economic growth rate of more than 8% for years, largely through the energy-intensive construction of infrastructure such as highways, railways and cities. Between 2005 and 2010, China increased its thermal-power generation by 63%, pig-iron and cement production by 74% and 76%, respectively, and vehicle production by 220% (ref. 5).

Although China has made tremendous efforts to limit air pollution, such as requiring coal-fired power plants to install flue-gas desulphurization systems and strengthening vehicle-emissions standards, these measures have not kept up with the growth of its economy and fossil-fuel use. We estimate that new equipment reduced SO_2 emissions from China's power plants by 1.5 million tonnes in 2005 and by 17.5 million tonnes in 2010 — 54% of the country's total SO_2 emissions in 2005 (32.3 million tonnes). But nationwide, total SO_2 emissions only decreased by 11% (to 28.7 million tonnes in 2010) because

those from other sectors grew (see 'China's emissions battle'). Coal usage rose by 44% (955 million tonnes), more than one-third of which was consumed by industrial facilities (such as iron, steel and cement works) that have no desulphurization systems.

The low priority given to environmental protection and the lack of cooperation among various government agencies also hampers air-quality control. China's Ministry of Environmental Protection (MEP) manages pollutant discharge, but it is a weak player within the government system. Its decisions are often obstructed for economic reasons. For instance, in late 2011, Chinese oil companies caused a delay in the planned 2012 implementation of a stricter vehicle-emissions standard (equivalent to Europe's Euro IV standard) because they were unable to provide the necessary low-sulphur oil. Any delay is a big strike against the environment, particularly as vehicle emissions continue to rise.

Such stories are rife in local governments, which like to promote heavy industry to stimulate regional economies. Local environmental agencies are often forced to back these projects just because they are affiliated to local governments. Yet, pushed by the public, the willingness and enthusiasm of China's government for curbing air pollution has never been so strong. It is a golden opportunity for

the nation to make a change — to free itself from the trap between economic development and environmental pollution.

CONSTRAIN COAL

Because China will continue to rely on fossil fuels for the next 20 years, the government should change its thinking. Instead of trying to use more energy to ensure economic growth regardless of the consequences, it should promote development with constrained fossil-fuel use. A cap could be set for national total coal consumption, and economic plans developed under this constraint. Otherwise, emissions from increased energy use will offset any gains from emissions control. Tertiary service industries and high-technology projects could be promoted instead of energy-intensive ones.

Greater authority should be given to environmental agencies at various levels of government. The MEP should be granted more power to implement its policies and enforce regulations. A vertical administrative structure would ensure that local environmental agencies report directly to the MEP.

The impact on the global environment should be considered when formulating China's air-quality strategies, and balanced plans should be developed for each pollution source. As well as reducing SO₂ emissions, the government should endorse measures to limit soot. Controlling emissions from diesel vehicles should be a priority, and oil companies should be brought into accordance with environmental standards. For future facilities that will control SO₂ and NO_x emissions, the government should equip them with specific technologies to remove mercury, such as activated carbon injection.

Addressing air pollution in China is a unique platform for researchers in atmospheric chemistry. Many scientific issues — such as secondary organic aerosol formation — remain to be explored. Practical control technologies for ultrafine particles and volatile organic compounds must be developed. Multinational collaboration is urgently needed; the government should make funds available to bring outstanding international scientists to China to help combat its air-pollution challenges. We all stand to benefit. ■

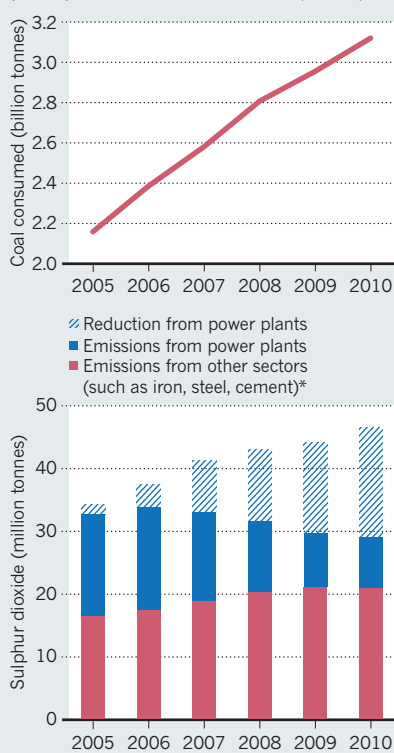
Qiang Zhang, Kebin He and Hong Huo

are professors in atmospheric chemistry, environmental science and energy systems at Tsinghua University, Beijing, China.
e-mail: hekb@tsinghua.edu.cn

1. Yang, F. *et al.* *Atmos. Chem. Phys.* **11**, 5207–5219 (2011).
2. Zhang, L. *et al.* *Atmos. Chem. Phys.* **8**, 6117–6136 (2008).
3. Shindell, D. *et al.* *Atmos. Chem. Phys.* **8**, 7101–7113 (2008).
4. Shindell, D. *et al.* *Nature Clim. Change* **1**, 59–66 (2011).
5. National Bureau of Statistics of China *China Statistical Yearbook* (2006; 2011).

CHINA'S EMISSIONS BATTLE

Rising coal use (top) has increased emissions of sulphur dioxide from some sectors, even though power plants emit less than in 2005 (bottom).





Q&A Michael Frayn

The playful dramatist

Author and playwright Michael Frayn explores the wellsprings of creativity through farce, philosophy and the history of science. His eclectic output ranges from non-fiction books such as *The Human Touch* (2006) to plays including *Noises Off* (1982) and *Copenhagen* (1998) — which explores the 1941 meeting between quantum physicists Niels Bohr and Werner Heisenberg, with Frayn imagining their discussions on the morality of working on nuclear weapons. With his latest novel, *Skios*, coming out next month, he talks about determinism and the paradox of existence.

What is *Skios* about?

It is a farce about an institute on a Greek island that has invited a lecturer to talk about the organization of science. The wrong lecturer shows up. I light-heartedly touch on determinism — the old question of whether human contributions to events are predetermined or whether they can't be understood in the context of causality. My view is the latter. There are two interesting things about farce. One is seeing human beings reduced to the level of machines, unable to control situations. The other is seeing people desperately thinking of ways to cope with difficult situations, inventing lies that they hope will get them out of the difficulties they're in, and of course making their difficulties worse.

Are you poking fun at the idea that human thought can be organized?

A tiny bit. People are endlessly surprised by the imagination. I'm struck by something that comes into *Copenhagen*: how physicists Otto Frisch and Rudolf Peierls came to understand how little fissile material you need to make a

nuclear chain reaction. It was because they were messing around. Everyone assumed that you would need tonnes, and there wasn't a chance of producing enough. One of the pair realized that fission had been observed only with slow neutrons, but we might want a formula that covered neutrons of all speeds. And he worked one out. Then the other one said, supposing we did have as much fissile material as we wanted, how much would be needed? And he applied the formula and discovered that it would be a matter of kilograms. These researchers weren't addressing the specific problem of building a bomb — they were working off the tops of their heads.

Why did you start writing about science?

I studied philosophy at university, and couldn't help but come across scientific questions, particularly in connection with



Skios

MICHAEL FRAYN
Faber & Faber/
Metropolitan Books:
2012. 224 pp.
£15.99/\$25

quantum mechanics. I had always had a faint idea of Heisenberg and Bohr's work, but never thought of writing about it until I read *Heisenberg's War* by Thomas Powers. Why did Heisenberg go to Copenhagen in 1941, and what were his motives in working on the German nuclear programme? Was he actually trying to build a bomb? Although there is obviously no parallel between this uncertainty and Heisenberg's uncertainty principle in quantum mechanics, there is a similar impossibility of ever reaching beyond a certain point. The result was *Copenhagen* — a play about epistemology that happens to be played out in terms of science.

Can drama teach science?

I don't think the theatre is a very good medium for explaining complex ideas. No one ignorant of nuclear physics would come out of *Copenhagen* thinking that they understood it.

In *The Human Touch*, you write about how the mind constrains our understanding of the world.

It is a paradox: we know perfectly well that we're irrelevant to the process of the ►

► **NATURE.COM**
For more on theatre
and the brain:
go.nature.com/eklgej

A. DALMAU/EPIC/CORBIS

► Universe — but there is nothing we can say about the Universe except in terms of what we see and think. I'm not suggesting that we make it all up arbitrarily. We're constrained by something, but it is extremely difficult to say what it is.

Some scientists would argue with that.

I can see how resistant scientists are to that side of the paradox. I was invited to CERN near Geneva, Switzerland, to talk about *The Human Touch*, and it was really daunting. They had appointed a jury that asked detailed questions. One of the jury members said beforehand: "We're going to haul you over the coals." It seemed to me — although they were all very charming and friendly about it — that they were unreconstructed Platonists. They believed that numbers and the laws of science are objective entities, whereas I think that they are constructs that we place on the world to understand it.

As a non-scientist, are you confident in writing about science?

Fortunately, professional science writers and scientists have made enormous efforts to get through to lay audiences. But people like the physicist Richard Feynman insist that if you haven't got mathematics you're never really going to understand physics — it is like trying to explain music to the tone-deaf. I made a lot of mistakes writing *Copenhagen*, in spite of getting the text read. I got letters from scientists pointing out basic errors. But I was struck by their generous tone.

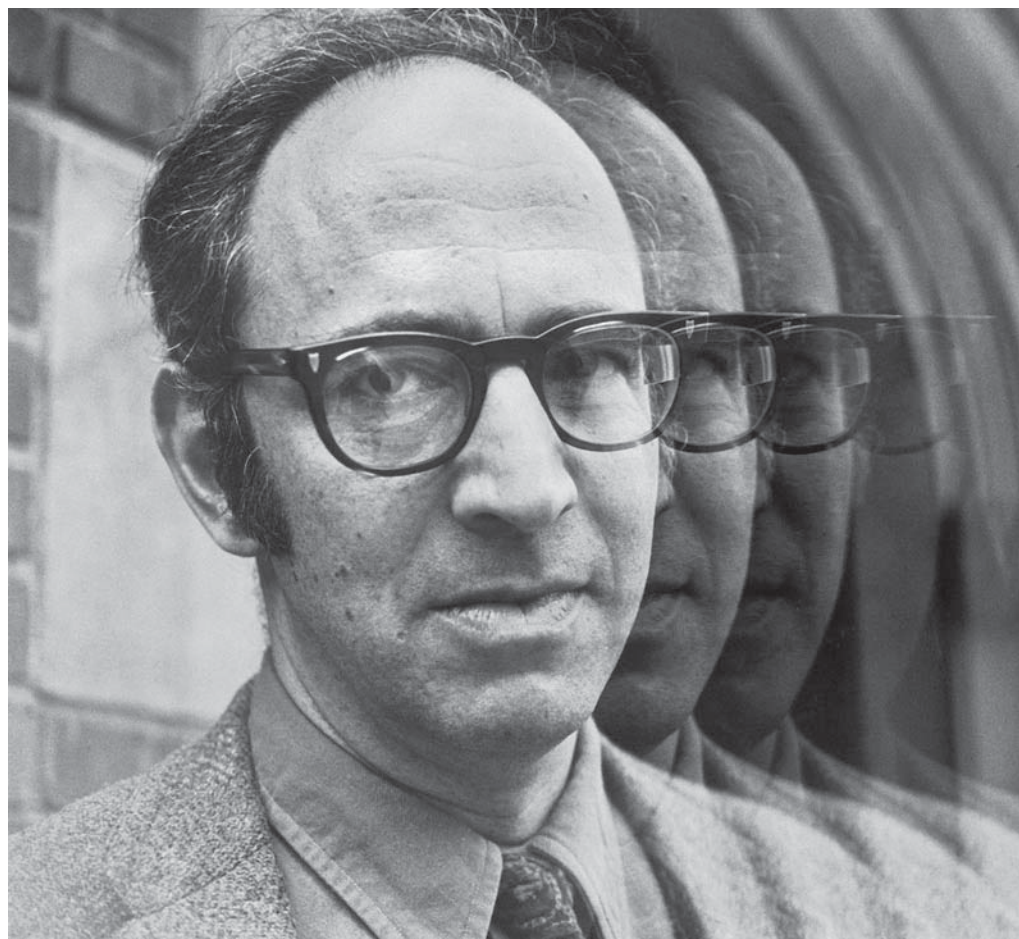
How do you approach writing?

As a writer, you can't think, "I'd like to write a play about stem-cell research and there will be these characters." It doesn't work like that: ideas just seem to fall into your head out of nowhere, and develop of their own accord. So there is resonance with the case of Peierls and Frisch, or the chemist August Kekulé dreaming about the structure of the benzene ring. There is an unconscious leap, a synthesis, that goes on, even though much science is about trying to find a specific answer to a specific problem.

So playwrights run experiments too?

Plays are called plays for a good reason. As a playwright, you are saying, what if we had enough uranium-235, or what if somebody discovered that their brother was their father, and you take over from these fictitious situations. It is messing around, but messing around often has serious results. ■

INTERVIEW BY RICHARD VAN NOORDEN



Thomas Kuhn recognized the importance of revolutionary changes, or 'paradigm shifts', in science.

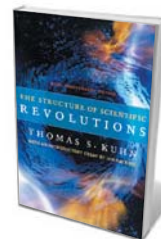
IN RETROSPECT

The Structure of Scientific Revolutions

David Kaiser marks the 50th anniversary of an exemplary account of the cycles of scientific progress.

Fifty years ago, a short book appeared under the intriguing title *The Structure of Scientific Revolutions*. Its author, Thomas Kuhn (1922–1996), had begun his academic life as a physicist but had migrated to the history and philosophy of science. His main argument in the book — his second work, following a study of the Copernican revolution in astronomy — was that scientific activity unfolds according to a repeating pattern, which we can discern by studying its history.

Kuhn was not at all confident about how *Structure* would be received. He had been



The Structure of Scientific Revolutions: 50th Anniversary Edition

THOMAS S. KUHN (WITH AN INTRODUCTION BY IAN HACKING)
Univ. Chicago Press: 2012.
264 pp. \$45, £29

denied tenure at Harvard University in Cambridge, Massachusetts, a few years before, and he wrote to several correspondents after the book was published that he felt he had stuck his neck "very far out". Within months,

B. PIERCE/TIME LIFE/GETTY

► Universe — but there is nothing we can say about the Universe except in terms of what we see and think. I'm not suggesting that we make it all up arbitrarily. We're constrained by something, but it is extremely difficult to say what it is.

Some scientists would argue with that.

I can see how resistant scientists are to that side of the paradox. I was invited to CERN near Geneva, Switzerland, to talk about *The Human Touch*, and it was really daunting. They had appointed a jury that asked detailed questions. One of the jury members said beforehand: "We're going to haul you over the coals." It seemed to me — although they were all very charming and friendly about it — that they were unreconstructed Platonists. They believed that numbers and the laws of science are objective entities, whereas I think that they are constructs that we place on the world to understand it.

As a non-scientist, are you confident in writing about science?

Fortunately, professional science writers and scientists have made enormous efforts to get through to lay audiences. But people like the physicist Richard Feynman insist that if you haven't got mathematics you're never really going to understand physics — it is like trying to explain music to the tone-deaf. I made a lot of mistakes writing *Copenhagen*, in spite of getting the text read. I got letters from scientists pointing out basic errors. But I was struck by their generous tone.

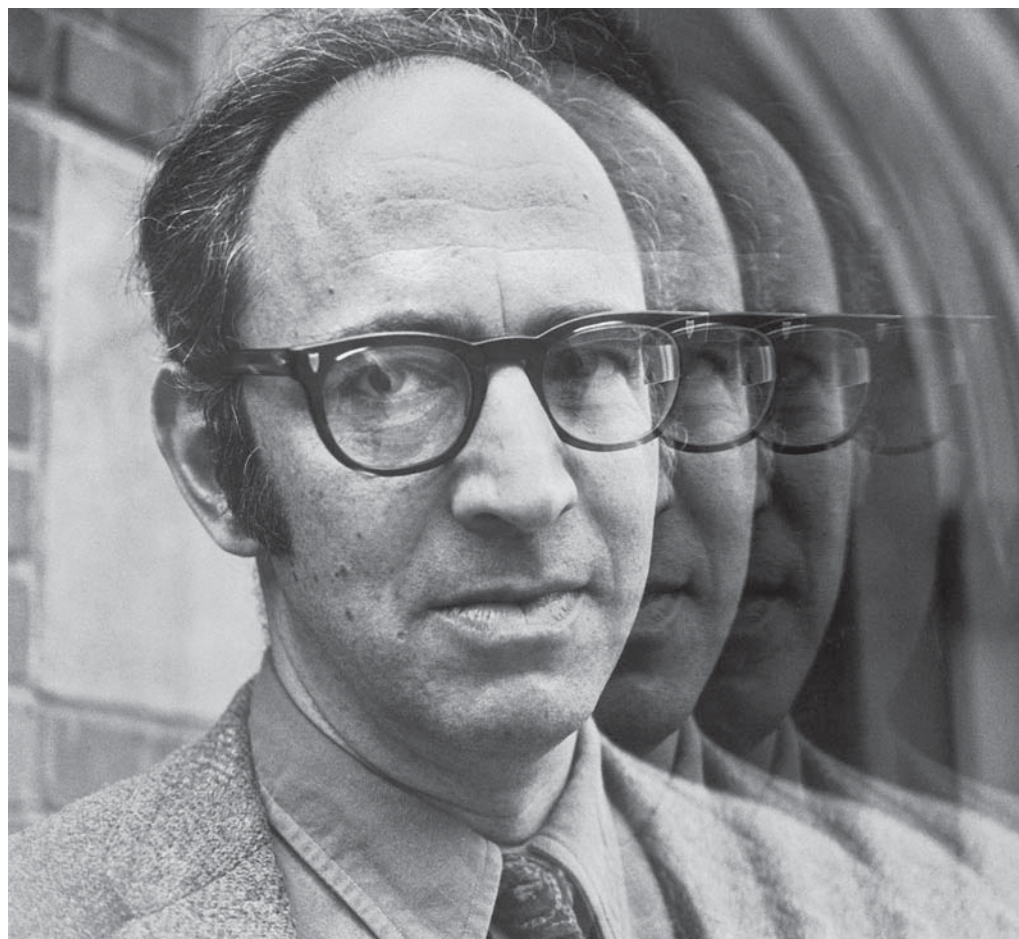
How do you approach writing?

As a writer, you can't think, "I'd like to write a play about stem-cell research and there will be these characters." It doesn't work like that: ideas just seem to fall into your head out of nowhere, and develop of their own accord. So there is resonance with the case of Peierls and Frisch, or the chemist August Kekulé dreaming about the structure of the benzene ring. There is an unconscious leap, a synthesis, that goes on, even though much science is about trying to find a specific answer to a specific problem.

So playwrights run experiments too?

Plays are called plays for a good reason. As a playwright, you are saying, what if we had enough uranium-235, or what if somebody discovered that their brother was their father, and you take over from these fictitious situations. It is messing around, but messing around often has serious results. ■

INTERVIEW BY RICHARD VAN NOORDEN



Thomas Kuhn recognized the importance of revolutionary changes, or 'paradigm shifts', in science.

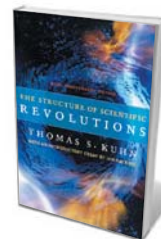
IN RETROSPECT

The Structure of Scientific Revolutions

David Kaiser marks the 50th anniversary of an exemplary account of the cycles of scientific progress.

Fifty years ago, a short book appeared under the intriguing title *The Structure of Scientific Revolutions*. Its author, Thomas Kuhn (1922–1996), had begun his academic life as a physicist but had migrated to the history and philosophy of science. His main argument in the book — his second work, following a study of the Copernican revolution in astronomy — was that scientific activity unfolds according to a repeating pattern, which we can discern by studying its history.

Kuhn was not at all confident about how *Structure* would be received. He had been



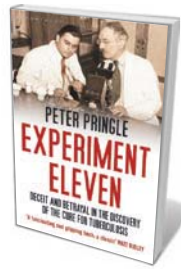
The Structure of Scientific Revolutions: 50th Anniversary Edition

THOMAS S. KUHN (WITH AN INTRODUCTION BY IAN HACKING)
Univ. Chicago Press: 2012.
264 pp. \$45, £29

denied tenure at Harvard University in Cambridge, Massachusetts, a few years before, and he wrote to several correspondents after the book was published that he felt he had stuck his neck "very far out". Within months,

B. PIERCE/TIME LIFE/GETTY

Books in brief



Experiment Eleven: Deceit and Betrayal in the Discovery of the Cure for Tuberculosis

Peter Pringle BLOOMSBURY 288 pp. £18.99 (2012)

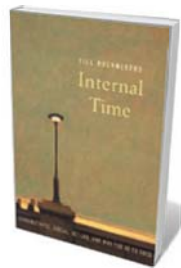
The 1943 discovery of a drug treatment for tuberculosis did much to kick-start big pharma. But this is a knotted tale, deftly unpicked by investigative journalist Peter Pringle. We learn that Albert Schatz, a US graduate student, found streptomycin in the eponymous 11th experiment on a farmyard bacterium — but that his research director, Selman Waksman, took the credit, along with patent royalties and a Nobel prize. A chance rediscovery brought Schatz the reputation he deserves.



The Forever Fix: Gene Therapy and the Boy Who Saved It

Ricki Lewis ST MARTIN'S 336 pp. \$25.99 (2012)

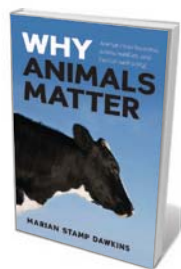
This popularized examination of gene therapy hinges on a breakthrough case: Corey Haas's recovery from Leber's congenital amaurosis type 2, which had made him virtually blind at the age of eight. Medical writer Ricki Lewis interweaves science, the history of medical trial and error, and human stories. The contrast can be intense, running from the death in 1999 of teenager Jesse Gelsinger, from a reaction to gene therapy intended to combat his liver disease, to radical successes in some children with adenosine deaminase deficiency.



Internal Time: Chronotypes, Social Jet Lag, and Why You're So Tired

Till Roenneberg HARVARD UNIVERSITY PRESS 288 pp. \$26.95 (2012)

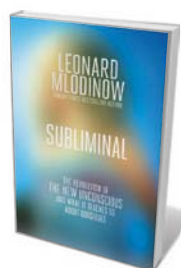
Time really is of the essence, says medical psychologist Till Roenneberg. By neglecting our body clocks — which rarely run in synchrony with the crazily cranked-up pace of modern life — we can develop “social jetlag”, endangering our health and careers. Roenneberg has built his book on decades of research in everything from fungi and single-celled organisms to humans. In brilliantly minimalist terms, he explains the temporal mismatches behind teen exhaustion, early birds and night owls, and sleep phobia.



Why Animals Matter: Animal consciousness, animal welfare, and human well-being

Marian Stamp Dawkins OXFORD UNIVERSITY PRESS 224 pp. £16.99 (2012)

Too little science and too much anthropomorphism have made our approaches to animal welfare a shambles, says ethologist Marian Stamp Dawkins. Her radical rethink involves linking their welfare with our own to harness a powerful driver of change: human self-interest. Dawkins advises sidestepping the question of animal consciousness to focus on animal health and hard-wired ‘wants’ such as foraging, to benefit both groups. Also key is never letting up on research into our intertwined existences, she says.



Subliminal: The Revolution of the New Unconscious and What it Teaches Us About Ourselves

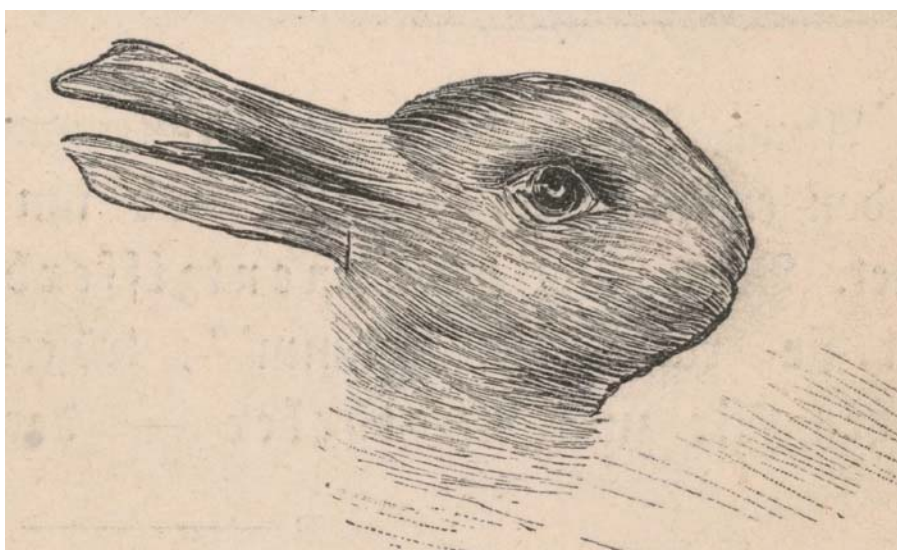
Leonard Mlodinow PANTHEON 272 pp. \$25.95 (2012)

Perception “below the threshold of consciousness”, as Carl Jung put it, is here pushed into the limelight. Physicist Leonard Mlodinow shows how humans have “parallel tiers” of a conscious brain superimposed on an unconscious mind. Drawing on research and anecdotes, Mlodinow explores the pattern-matching, gap-filling role of the unconscious in perception, memory, sociality, emotions and self-estimation. An illuminating journey through a hidden world.

however, some people were proclaiming a new era in the understanding of science. One biologist joked that all commentary could now be dated with precision: his own efforts had appeared “in the year 2 B.K.”, before Kuhn. A decade later, Kuhn was so inundated with correspondence about the book that he despaired of ever again getting any work done.

By the mid-1980s, *Structure* had achieved blockbuster status. Nearly a million copies had been sold and more than a dozen foreign-language editions published. The book became the most-cited academic work in all of the humanities and social sciences between 1976 and 83 — cited more often than classic works by Sigmund Freud, Ludwig Wittgenstein, Noam Chomsky, Michel Foucault or Jacques Derrida. The book was required reading for undergraduates in classes across the curriculum, from history and philosophy to sociology, economics, political science and the natural sciences. Before long, Kuhn's phrase “paradigm shift” was showing up everywhere from business manuals to cartoons in *The New Yorker*.

Kuhn began thinking about his project 15 years before it was published, while he was working on his doctorate in theoretical physics at Harvard. He became interested in ►



The duck–rabbit figure shows how two pictures can be derived from the same evidence.

► developmental psychology, avidly reading works by Swiss psychologist Jean Piaget about the stages of cognitive development in children.

Kuhn saw similar developmental stages in entire sciences. First, he said, a field of study matures by forming a paradigm — a set of guiding concepts, theories and methods on which most members of the relevant community agree. There follows a period of “normal science”, during which researchers further articulate what the paradigm might imply for specific situations.

In the course of that work, anomalies necessarily arise — findings that differ from expectations. Kuhn had in mind episodes such as the accidental discoveries of X-rays in the late nineteenth century and nuclear fission in the early twentieth. Often, Kuhn argued, the anomalies are brushed aside or left as problems for future research. But once enough anomalies have accumulated, and all efforts to assimilate them to the paradigm have met with frustration, the field enters a state of crisis. Resolution comes only with a revolution, and the inauguration of a new paradigm that can address the anomalies. Then the whole process repeats with a new phase of normal science. Kuhn was especially struck by the cyclic nature of the process, which ran counter to then-conventional ideas about scientific progress.

At the heart of Kuhn’s account stood the tricky notion of the paradigm. British philosopher Margaret Masterman famously isolated 21 distinct ways in which Kuhn used the slippery term throughout his slim volume. Even Kuhn himself came to realize that he had saddled the word with too much baggage: in later essays, he separated

► **NATURE.COM**
For more on the
social practices of
science:
go.nature.com/ghydti

his intended meanings into two clusters. One sense referred to a scientific community’s reigning theories and methods. The second meaning, which Kuhn argued was both more original and more important, referred to exemplars or model problems, the worked examples on which students and young scientists cut their teeth. As Kuhn appreciated from his own physics training, scientists learned by immersive apprenticeship; they had to hone what Hungarian chemist and

“Scientists have no way to compare concepts on either side of a scientific revolution.”

philosopher of science Michael Polanyi had called “tacit knowledge” by working through large collections of exemplars rather than by memorizing explicit rules or theorems. More than

most scholars of his era, Kuhn taught historians and philosophers to view science as practice rather than syllogism. Most controversial was Kuhn’s claim that scientists have no way to compare concepts on either side of a scientific revolution. For example, the idea of ‘mass’ in the Newtonian paradigm is not the same as in the Einsteinian one, Kuhn argued; each concept draws meaning from separate webs of ideas, practices and results. If scientific concepts are bound up in specific ways of viewing the world, like a person who sees only one aspect of a Gestalt psychologist’s duck–rabbit figure, then how is it possible to compare one concept to another? To Kuhn, the concepts were incommensurable: no common measure could be found with which to relate them, because scientists, he argued, always interrogate nature through a given paradigm.

Perhaps the most radical thrust of Kuhn’s analysis, then, was that science might not be progressing toward a truer representation of

the world, but might simply be moving away from previous representations. Knowledge need not be cumulative: when paradigms change, whole sets of questions and answers get dropped as irrelevant, rather than incorporated into the new era of normal science. In the closing pages of his original edition, Kuhn adopted the metaphor of Darwinian natural selection: scientific knowledge surely changes over time, but does not necessarily march towards an ultimate goal.

And so, 50 years later, we are left with our own anomaly. How did an academic book on the history and philosophy of science become a cultural icon? *Structure* was composed as an extended essay rather than a formal monograph: it was written as an entry on the history of science for the soon-to-be-defunct *International Encyclopedia of Unified Science*. Kuhn never intended it to be definitive. He often described the book (even in its original preface) as a first pass at material that he intended to address in more detail later.

To me, the book has the feel of a physicist’s toy model: an intentionally stripped-down and simplified schematic — an exemplar — intended to capture important phenomena. The thought-provoking thesis is argued with earnestness and clarity, not weighed down with jargon or lumbering footnotes. The more controversial claims are often advanced in a suggestive rather than declarative mode. Perhaps most important, the book is short: it can be read comfortably in a single sitting.

For the 50th-anniversary edition, the University of Chicago Press has included an introductory essay by renowned Canadian philosopher Ian Hacking. Like Kuhn, Hacking has a gift for clear exposition. His introduction provides a helpful guide to some of the thornier philosophical issues, and gives hints as to how historians and philosophers of science have parted with Kuhn.

The field of science studies has changed markedly since 1962. Few philosophers still subscribe to radical incommensurability; many historians focus on sociological or cultural features that received no play in Kuhn’s work; and topics in the life sciences now dominate, whereas Kuhn focused closely on physics. Nevertheless, we may still admire Kuhn’s dexterity in broaching challenging ideas with a fascinating mix of examples from psychology, history, philosophy and beyond. We need hardly agree with each of Kuhn’s propositions to enjoy — and benefit from — this classic book. ■

David Kaiser is Germeshausen Professor of the History of Science at the Massachusetts Institute of Technology in Cambridge. His latest book is *How the Hippies Saved Physics* (Norton, 2011).
e-mail: dikaiser@mit.edu

Correspondence

Primate studies: hear the public's views

A painful irony in the disrupted flow of primates to US research labs (*Nature* **483**, 381–382; 2012) is that the number being used in laboratory experiments is at an all-time high.

If the scientific community is to maintain the support and trust of the public, which funds much of its work, then research practices and policies should change to reflect society's views on what constitutes the ethical treatment of animals. These changes need to be speeded up.

Committees that review and approve animal experiments at US facilities should not be dominated by those who work in animal labs and have vested interests in continuing animal research (L. A. Hansen *et al. Animals* **2**, 68–75; 2012). There are too few members of the public on these US committees, and those who are involved say that they are often marginalized. In other countries, such as Sweden and Australia, half or one-third of committees must comprise non-scientists and animal-welfare representatives.

If scientists continue to disregard the substantial and growing public opposition to harmful research on primates and other animals, more protest campaigns are inevitable.

Lawrence A. Hansen *University of California, San Diego, La Jolla, California, USA.*
lahansen@ucsd.edu

Primate studies: fix welfare issues first

Record numbers of non-human primates are being used in US labs, so it is unlikely that limiting imports will hold back vital areas of research as you imply (*Nature* **483**, 381–382; 2012).

A report from the American Anti-Vivisection Society (AAVS; see go.nature.com/gbqlle) indicates that imports of monkeys born to wild-caught

parents quadrupled during 1998–2008. Conservationists are concerned about global trade in crab-eating macaques (*Macaca fascicularis*), the import of which has doubled in recent years.

Scientists must urgently address the extreme animal-welfare issues surrounding these imports. The AAVS report, which is based on information from US federal agencies and scientific studies, has revealed that monkeys destined for US labs typically endure long, gruelling air and land transportation; entire groups have been killed after quarantine on testing positive for tuberculosis; many die from transport injuries or stress in quarantine; and survivors show negative physiological and behavioural effects for several months after the journey.

More airlines are likely to back away from a dirty job that they are ill-equipped to do properly. **Crystal Miller-Spiegel** *American Anti-Vivisection Society, Jenkintown, Pennsylvania, USA.*
cmillerspiegel@aavs.org

Primate studies: trials don't always translate

In your discussion on the campaign against animal research (*Nature* **483**, 373–374; 2012), you mention a study in macaques that has moved into early clinical trials in humans, with promising results. Sadly, there is a yawning chasm between early promise in trials and efficacy.

The US Food and Drug Administration reports that more than 90% of trials fail (see go.nature.com/h2365q), even though the treatments tested, by definition, were considered sufficiently efficacious and safe in animals to merit a clinical-trial licence.

Many other treatments to protect the brain after stroke have failed in humans, despite success in rodent and primate trials (V. E. Collins *et al. Ann. Neurol.* **59**, 467–477; 2006). None of the 85 or so candidate HIV vaccines

that were effective in primates has so far worked in humans (J. Bailey *Altern. Lab. Anim.* **36**, 381–428; 2008). The monoclonal antibody that caused severe inflammatory reactions in a 2006 clinical trial at Northwick Park Hospital, London, caused no problems in primates at 500 times the dose given to the human volunteers.

The public is rightly concerned about the transportation of primates for questionable experimental purposes. These cannot justify the degree of suffering involved during capture, in breeding and holding facilities and during lengthy transportation (see go.nature.com/svbuvi).

Michelle Thew *British Union for the Abolition of Vivisection, London, UK.*
michelle.thew@buav.org

Higgs can claim name of massive boson

Attempts to rule against naming the Higgs boson after physicist Peter Higgs suggest that political correctness is taking over from scholarship (*Nature* **483**, 374; 2012). Your suggestion that the name Higgs should be retained for reasons akin to business branding is hardly better. Higgs has a unique claim to the massive boson in question.

My book *The Infinity Puzzle* (Oxford Univ. Press, 2011) covers the history of the Higgs hypothesis in detail. It is true that Higgs is one of several theorists who, in 1964, independently discovered how to give mass to fundamental particles, and that it would be inappropriate to refer to the hypothesis of mass generation as the 'Higgs mechanism'. However, it was Higgs alone who drew attention to the massive boson whose detection can prove the hypothesis. So naming the boson after him, as Ben Lee did in 1972, is justifiable.

Frank Close *University of Oxford, Oxford, UK.*
f.close1@physics.ox.ac.uk

More credit due to India's scientists

Any increase in India's science budget for 2012–13 is likely to be wiped out by a 5–10% rise in the cost of research commodities, owing to the country's high rate of inflation (*Nature* **483**, 384; 2012). Neither will the modest extra funding tackle the glut of unemployed science PhD graduates (for example, around 60% of female science PhDs do not have a research position).

The reasons for this situation are not just economic. In my opinion, India's policy-makers are failing to recognize scientists' achievements. In May last year, for example, environment minister Jairam Ramesh intimated that India's elite institutions, which include the Indian Institutes of Technology and of Management, fall short of world-class standards; the head of the prime minister's scientific advisory council, C. N. R. Rao, seems to agree with this view (see go.nature.com/snnbtt).

Yet India is ranked 11th in the world by number of publications and 16th on the basis of total citations during the past 10 years. Many of these publications were in leading international journals. Instead of squandering this talent, the government should provide the incentive and the means for the nation to fulfil its potential.

Jagadeesh Bayry *Institut National de la Santé et de la Recherche Médicale, Paris, France.*
jagadeesh.bayry@crc.jussieu.fr

CONTRIBUTIONS

Correspondence may be sent to correspondence@nature.com after consulting the author guidelines at go.nature.com/cmchno. Alternatively, readers may comment online: www.nature.com/nature.

F. Sherwood Rowland

(1927–2012)

Atmospheric chemist who linked human activity to ozone depletion.

Chlorofluorocarbons (CFCs) were a triumph of the chemical industry and a mere curiosity in atmospheric science when Sherwood (Sherry) Rowland, with his postdoc Mario Molina, recognized in 1973 that these seemingly inert gases posed a threat to Earth's ozone layer. Returning home one evening, Rowland remarked to his wife Joan that the research "is going very well, but it may mean the end of the world".

In their laboratory at the University of California, Irvine, Molina and Rowland had discovered that CFC-11 (CFCl_3) and CFC-12 (CF_2Cl_2), then widely used as refrigerants and aerosol propellants, readily absorbed ultraviolet light and broke down to release reactive chlorine. This work was the first step in tracing the causal chain linking industrial production of CFCs with global ozone depletion — and won Rowland and Molina the 1995 Nobel Prize in Chemistry, shared with Dutch chemist Paul Crutzen.

Surrounded by his family at his home in Corona del Mar, California, Rowland died on 10 March, aged 84, from complications of Parkinson's disease. He was born in Delaware, Ohio; his mother was a Latin teacher and his father taught mathematics at Ohio Wesleyan University in Delaware, where Rowland attended college after graduating from high school at 15. When he was old enough, he enlisted in the US Navy. As a lanky athlete, he readily found a home in sports teams in the Navy and later in graduate school at the University of Chicago, Illinois, where he played baseball for the university and for a semi-professional team.

Rowland earned his PhD in nuclear chemistry at Chicago under chemist Willard Libby and was taught by four other faculty members; counting Sherry, all six would later receive Nobel prizes. He met Joan at Chicago, and they moved to take up his early jobs at Princeton University in New Jersey and at the University of Kansas in Lawrence. In 1964, Rowland accepted an offer to start up the chemistry department at a new University of California campus in the then-unbuilt city of Irvine. Later, with atmospheric chemist Ralph Cicerone, he also helped to found the Earth system science department.

The elegance of Molina and Rowland's 1974 *Nature* paper remains impressive to today's atmospheric chemists, who live in a world of satellite observations and supercomputers. Stratospheric chemistry at the time was based on balloon-borne samplers of

trace gases and on one-dimensional models that could now easily run on a smart phone.

Nonetheless, the pair measured the ultraviolet cross sections of CFC-11 and -12 in the lab, calculated their photolytic destruction rates in the atmosphere and derived their atmospheric lifetimes as 50–100 years. They reviewed industrial production and emission of CFCs, projected the build-up and release of chlorine atoms in the stratosphere and concluded that ozone depletion



was likely and would be long-lived, even with remediation. This work has been borne out, in detail, by nearly four decades of research.

Rowland and Molina's work started an environmental movement that began with scientists, led by Rowland, urging the elimination of CFCs. It remains the best success story for global cooperation on a worldwide environmental threat. The activism led to the 1978 ban by the US Environmental Protection Agency on CFC use in aerosol cans, and finally in 1990 to the complete phasing out of CFC production by the Montreal Protocol and its amendments.

In his unwillingness to back down from the implications of his work, Rowland became a role model to many of us, and remains so. This was a threat to some — particularly the CFC industry, but also, less understandably, to some scientific colleagues. For many years, Rowland experienced personal threats as well as irrational attacks on the science.

Rowland's science always stood tall, as did he, and seemed inerrant. He kept up his

interests in ozone and environmental policy, but his research moved on. Soon after the ozone hole was discovered over Antarctica, he made major contributions with his graduate student Neil Harris to the detection of ozone depletion over the Northern Hemisphere. This work was crucial in persuading DuPont and other chemical companies to abandon CFCs in favour of hydrochlorofluorocarbons, which are less damaging to the ozone layer.

In the late 1970s, Rowland initiated a programme to monitor background concentrations of various gases, and that continues today. Six of its group members were working in the field when he passed away. His curiosity demanded an objective approach, and so it was when, working with his former student and then fellow professor (D.R.B.), he identified in 1995 that the unusual mix of high ozone and hydrocarbons in Mexico City was due to leaking propane stoves and heaters, rather than traffic. In 2011, he was involved in discussions regarding the mix of atmospheric hydrocarbons resulting from the Deepwater Horizon oil spill in the Gulf of Mexico.

Over almost five decades, Rowland was active in his research lab as well as teaching, playing tennis and having collegial discussions over lunch. When not travelling, he could be seen carrying his briefcase in one hand, with a pile of papers under the other arm, to and from his office. He was a prolific note-taker, filling a notebook in a week. This practice intimidated one of us (M.J.P.), who, while giving a talk at an international conference, first encountered Sherry in the front row, taking assiduous notes and then asking a terrifying, brutal, yet constructive question.

Rowland treated everyone like a colleague. He disarmingly considered questions from any listener with the depth and profundity due a scientific peer. This trait was appreciated by students, friends and family. To Sherry, the question was of foremost importance; it was at the core of his scientific quest. His passing ended a unique career that merged chemistry and atmospheric sciences, leading to a new partnership between science and policy for the protection of our planet. ■

Michael J. Prather is Fred Kavli Chair in Earth system science and **Donald R. Blake** is professor of chemistry at the University of California, Irvine, California, USA.
e-mail: mprather@uci.edu

C. PUMAVIC IRVINE

CLIMATE SCIENCE

Aerosols and Atlantic aberrations

A cutting-edge global climate model links atmospheric aerosol emissions to temperature variability in the North Atlantic Ocean, suggesting that human activity influences extreme weather events. [SEE LETTER P.228](#)

AMATO EVAN

Over the past century, the surface of the North Atlantic Ocean has gone through warm and cool periods that are not observed in other ocean basins. This Atlantic Multidecadal Oscillation (AMO)¹ is thought to affect climate processes² ranging from the current high levels of Atlantic hurricane activity to the devastating sub-Saharan droughts of the early 1980s. Although the influence of the AMO on extreme weather events has long been recognized, the physical processes underlying these temperature changes are not understood. On page 228 of this issue, Booth *et al.*³ report their use of a state-of-the-art model of Earth's climate to demonstrate that, at least over the past century, the AMO is largely the response of the upper ocean to changes in the concentration of pollution aerosols in the atmosphere. If correct, their results imply that the influence of human activity on the Atlantic regional climate is more pervasive than previously thought*.

The AMO is best depicted as the difference between average ocean surface temperatures over the North Atlantic and those over the global oceans⁴ (Fig. 1). It therefore reflects the deviation of the North Atlantic Ocean from global mean temperatures, which are dominated by the long-term warming that is forced by greenhouse gases. Conventional wisdom has held that the AMO is the natural result of internal processes in the Atlantic Ocean — most notably, fluctuations in deep-ocean circulation, as supported by multi-century climate-model studies⁵.

Booth *et al.*³ simulated the climate of the past 150 years using a version of a well-known climate model⁶ that includes up-to-date parameterizations of aerosol emissions, aerosol chemistry and interactions of aerosols and clouds. Nearly all of the observed decadal variability in North Atlantic surface temperatures was reproduced in their simulation, including

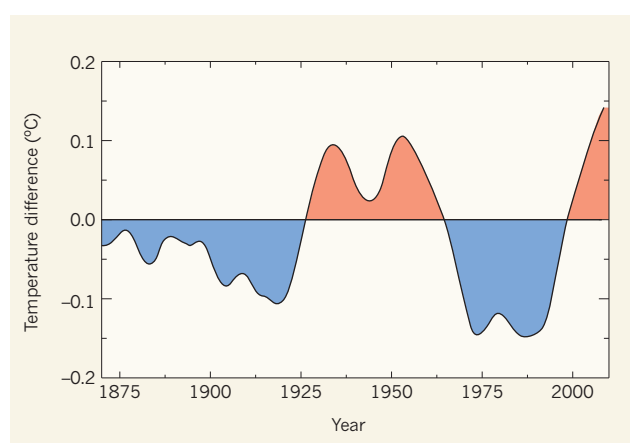


Figure 1 | The Atlantic Multidecadal Oscillation. The difference between average ocean surface temperatures over the North Atlantic and those over the global oceans has oscillated between cool and warm phases over the past hundred years or so, as depicted here. Booth *et al.*³ report that simulations of global climate link this temperature variability to atmospheric aerosol emissions.

the AMO and the long-term warming associated with increasing amounts of greenhouse gases. This is the first time that changes in sea surface temperatures have been reproduced to this degree of accuracy by a climate model. The authors' analysis of the model's output reveals that this variability about the upward temperature trend results from cooling associated with periodic volcanic eruptions, and from the build-up of polluting aerosols in the atmosphere that occurred from pre-industrial times until the late 1960s and early 1970s, when clean-air legislation in the United States and Europe was implemented.

So how do aerosols affect sea surface temperatures? When suspended over water, aerosols tend to cool the surface by increasing the local albedo (the ability to reflect sunlight), a phenomenon known as the aerosol direct effect. Aerosols caused by human activity may also act as nuclei around which water vapour in the atmosphere can condense. When more of these nuclei are available for condensation within a cloud, the number of water droplets in the cloud goes up and the average size of the water droplets goes down, making the cloud brighter so that it reflects more sunlight back out to space. This process is known as the cloud albedo effect, or the first aerosol indirect

effect. It is these aerosol–cloud interactions that have the most influence over the AMO in Booth and colleagues' model.

The idea that variations in aerosol concentration have caused decadal-scale changes in surface temperatures in the North Atlantic is not new. A body of work has emerged suggesting that elements of the AMO are externally forced by aerosols through direct and first indirect effects, and has implicated aerosols from volcanic eruptions⁷, West African dust storms⁸ and human activity^{4,9}. However, before Booth and colleagues' work, no study had incorporated the direct and indirect forcings from these various aerosol types to paint a coherent picture of temperature changes in the Atlantic Ocean that was consistent with both the observed

temporal variability and the dominant spatial structure of the changes.

Booth and colleagues' evidence³ that the AMO is caused by changes in the regional abundance of aerosols is compelling, but their results are sensitive to model parameterizations of microphysical processes, particularly the interaction between cloud water droplets and aerosols, that are not well constrained by observations. In addition, their model was unable to reproduce observed multidecadal variability in outbreaks of African dust storms¹⁰, which alter the temperature of the tropical Atlantic⁸; this may explain why the model does a poorer job of simulating temperatures in the tropical North Atlantic Ocean than it does in the extratropical regions. Furthermore, the authors' conclusion that internal variability of the Atlantic Ocean had a negligible role in shaping the AMO during the twentieth century is at odds with the findings of several previous studies^{5,11}. The reason for this discrepancy is not clear.

If Booth and colleagues' results³ can be corroborated, then they suggest that multidecadal temperature fluctuations of the North Atlantic are dominated by human activity, with natural variability taking a secondary role. This has many implications. Foremost among them is

*This article and the paper under discussion³ were published online on 4 April 2012.

that the AMO does not exist, in the sense that the temperature variations concerned are neither intrinsically oscillatory nor purely multidecadal.

Another implication concerns hurricanes. As noted earlier, quiescent and active periods of Atlantic hurricane activity have been linked² to the AMO. These swings in hurricane frequency and intensity might therefore be the regional response to variations in the concentration of pollutant aerosols against a background of global warming, and thus completely man-made. Similarly, human activity might have caused periods of drought within the Sudano-Sahel region of Africa and in northeastern Brazil.

As we try to predict the climate in a warming world, an increasing body of work suggests that aerosols may have regional effects as great as those caused by the global increase in atmospheric carbon dioxide. Booth and colleagues' work³ underscores the importance of understanding the diverse pathways by which humans alter the climate. ■

Amato Evan is in the Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia 22904, USA. e-mail: ate9c@virginia.edu

1. Kerr, R. A. *Science* **288**, 1984–1985 (2000).
2. Knight, J. R., Folland, C. K. & Scaife, A. *Geophys. Res. Lett.* **33**, L17706 (2006).

3. Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T. & Bellouin, N. *Nature* **484**, 228–232 (2012).
4. Mann, M. E. & Emanuel, K. A. *Eos* **87**, 233–244 (2006).
5. Delworth, T. L. & Mann, M. E. *Clim. Dyn.* **16**, 661–676 (2000).
6. Collins, W. J. *et al. Geosci. Model Dev.* **4**, 1051–1075 (2011).
7. Otterå, O. H., Bentsen, M., Drange, M. & Suo, L. *Nature Geosci.* **3**, 688–694 (2010).
8. Evan, A. T., Vimont, D. J., Heidinger, A. K., Kossin, J. P. & Bennartz, R. *Science* **324**, 778–781 (2009).
9. Chang, C.-Y., Chiang, J. C. H., Wehner, M. F., Friedman, A. R. & Ruedy, R. J. *J. Clim.* **24**, 2540–2555 (2011).
10. Prospero, J. M. & Lamb, P. J. *Science* **302**, 1024–1027 (2003).
11. Deser, C., Alexander, M. A., Xie, S.-P. & Phillips, A. S. *Annu. Rev. Mar. Sci.* **2**, 115–143 (2010).

GENE EXPRESSION

Running to stand still

Transcription factors regulate the expression of genes by binding to certain DNA sequences. But the outcome can be markedly different, depending on whether the binding is stable or short-lived. SEE LETTER P.251

TOMMY KAPLAN & NIR FRIEDMAN

To reproduce, differentiate or even just respond to changes in their surroundings, cells need to control the expression of thousands of genes. One way of doing this is to use transcription factors — proteins that bind to regulatory regions on their target genes and either activate or repress the transcription of DNA into RNA. Over the past decade, researchers have analysed the binding sites of hundreds of these proteins on the genomes of many organisms and cell types, and measured the gene-expression patterns within the same cells. In such studies, the overall degree of occupancy by a transcription factor at a regulatory region is commonly interpreted as an indication of the protein's ability to control the expression of the gene. However, transcription factors also bind to thousands of genes in a weak, and probably non-functional, manner¹. On page 251 of this issue, Lickwar *et al.*² illuminate this matter by reporting the results of a systematic, genome-wide study of the binding dynamics of a particular transcription factor. The authors find that transcription levels have a stronger link to the kinetics of binding than to the total occupancy of the factor.

The DNA-binding sites of the transcription factor Rap1 along the genome of the yeast *Saccharomyces cerevisiae* were mapped more than a decade ago³. The mapping used a genome-wide protein–DNA binding assay, known as chromatin immunoprecipitation (ChIP)-on-chip, or ChIP-seq, which identifies the genomic locations of a transcription factor over a huge number of live cells and therefore

averages the transcription-factor occupancy over a large population. This is still the method of choice in similar genome-wide studies. However, a high occupancy of a transcription factor at a specific site — as detected by this technique — can mean either that the factor is constantly bound to this DNA location in

some of the cells, or that it is transiently bound in many cells.

To distinguish between the two possibilities, Lickwar *et al.*² adapted a strategy, previously used to measure the turnover of DNA-bound proteins^{4–6}, to address the question of transcription-factor binding stability. The authors created yeast cells that produced two Rap1 variants, Rap1–Flag and Rap1–Myc, each one with a 'tag' consisting of a specific peptide that could be recognized by antibodies. Furthermore, Rap1–Flag was produced constantly, whereas Rap1–Myc's expression was experimentally inducible. The authors then measured the binding of each Rap1 variant to the yeast genome in a dense time series after Rap1–Myc induction. Although the inducible protein quickly outcompeted Rap1–Flag at

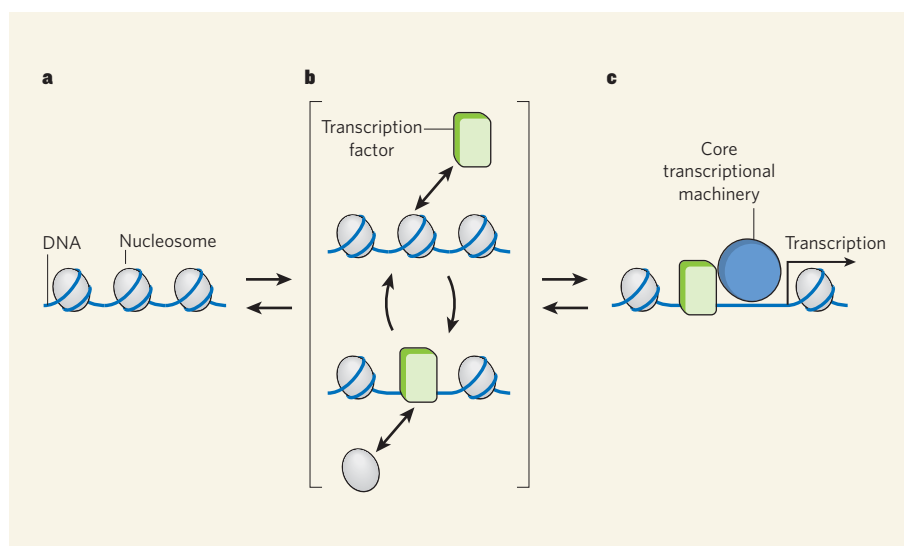


Figure 1 | Well-balanced gene expression. Transcription factors can activate the expression of genes by binding to certain regulatory regions on the genome. Lickwar *et al.*² studied, at the genomic scale, the binding dynamics of one of these proteins, and propose the following model for transcription-factor function. **a**, Most of the genome is wrapped around histone proteins to form nucleosomes, the basic units of DNA packaging into chromatin. **b**, Nucleosomes and transcription factors compete for binding to some regulatory regions. Transcription-factor binding to these regions occurs in short pulses, which are not sufficient for efficient transcription of the gene into RNA. **c**, When the transcription factor binds to its target site for longer periods, it recruits the core transcriptional machinery required to start transcription, leading to high transcription rates.

that the AMO does not exist, in the sense that the temperature variations concerned are neither intrinsically oscillatory nor purely multidecadal.

Another implication concerns hurricanes. As noted earlier, quiescent and active periods of Atlantic hurricane activity have been linked² to the AMO. These swings in hurricane frequency and intensity might therefore be the regional response to variations in the concentration of pollutant aerosols against a background of global warming, and thus completely man-made. Similarly, human activity might have caused periods of drought within the Sudano-Sahel region of Africa and in northeastern Brazil.

As we try to predict the climate in a warming world, an increasing body of work suggests that aerosols may have regional effects as great as those caused by the global increase in atmospheric carbon dioxide. Booth and colleagues' work³ underscores the importance of understanding the diverse pathways by which humans alter the climate. ■

Amato Evan is in the Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia 22904, USA. e-mail: ate9c@virginia.edu

1. Kerr, R. A. *Science* **288**, 1984–1985 (2000).
2. Knight, J. R., Folland, C. K. & Scaife, A. *Geophys. Res. Lett.* **33**, L17706 (2006).

3. Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T. & Bellouin, N. *Nature* **484**, 228–232 (2012).
4. Mann, M. E. & Emanuel, K. A. *Eos* **87**, 233–244 (2006).
5. Delworth, T. L. & Mann, M. E. *Clim. Dyn.* **16**, 661–676 (2000).
6. Collins, W. J. *et al. Geosci. Model Dev.* **4**, 1051–1075 (2011).
7. Otterå, O. H., Bentsen, M., Drange, M. & Suo, L. *Nature Geosci.* **3**, 688–694 (2010).
8. Evan, A. T., Vimont, D. J., Heidinger, A. K., Kossin, J. P. & Bennartz, R. *Science* **324**, 778–781 (2009).
9. Chang, C.-Y., Chiang, J. C. H., Wehner, M. F., Friedman, A. R. & Ruedy, R. J. *J. Clim.* **24**, 2540–2555 (2011).
10. Prospero, J. M. & Lamb, P. J. *Science* **302**, 1024–1027 (2003).
11. Deser, C., Alexander, M. A., Xie, S.-P. & Phillips, A. S. *Annu. Rev. Mar. Sci.* **2**, 115–143 (2010).

GENE EXPRESSION

Running to stand still

Transcription factors regulate the expression of genes by binding to certain DNA sequences. But the outcome can be markedly different, depending on whether the binding is stable or short-lived. SEE LETTER P.251

TOMMY KAPLAN & NIR FRIEDMAN

To reproduce, differentiate or even just respond to changes in their surroundings, cells need to control the expression of thousands of genes. One way of doing this is to use transcription factors — proteins that bind to regulatory regions on their target genes and either activate or repress the transcription of DNA into RNA. Over the past decade, researchers have analysed the binding sites of hundreds of these proteins on the genomes of many organisms and cell types, and measured the gene-expression patterns within the same cells. In such studies, the overall degree of occupancy by a transcription factor at a regulatory region is commonly interpreted as an indication of the protein's ability to control the expression of the gene. However, transcription factors also bind to thousands of genes in a weak, and probably non-functional, manner¹. On page 251 of this issue, Lickwar *et al.*² illuminate this matter by reporting the results of a systematic, genome-wide study of the binding dynamics of a particular transcription factor. The authors find that transcription levels have a stronger link to the kinetics of binding than to the total occupancy of the factor.

The DNA-binding sites of the transcription factor Rap1 along the genome of the yeast *Saccharomyces cerevisiae* were mapped more than a decade ago³. The mapping used a genome-wide protein–DNA binding assay, known as chromatin immunoprecipitation (ChIP)-on-chip, or ChIP-seq, which identifies the genomic locations of a transcription factor over a huge number of live cells and therefore

averages the transcription-factor occupancy over a large population. This is still the method of choice in similar genome-wide studies. However, a high occupancy of a transcription factor at a specific site — as detected by this technique — can mean either that the factor is constantly bound to this DNA location in

some of the cells, or that it is transiently bound in many cells.

To distinguish between the two possibilities, Lickwar *et al.*² adapted a strategy, previously used to measure the turnover of DNA-bound proteins^{4–6}, to address the question of transcription-factor binding stability. The authors created yeast cells that produced two Rap1 variants, Rap1–Flag and Rap1–Myc, each one with a 'tag' consisting of a specific peptide that could be recognized by antibodies. Furthermore, Rap1–Flag was produced constantly, whereas Rap1–Myc's expression was experimentally inducible. The authors then measured the binding of each Rap1 variant to the yeast genome in a dense time series after Rap1–Myc induction. Although the inducible protein quickly outcompeted Rap1–Flag at

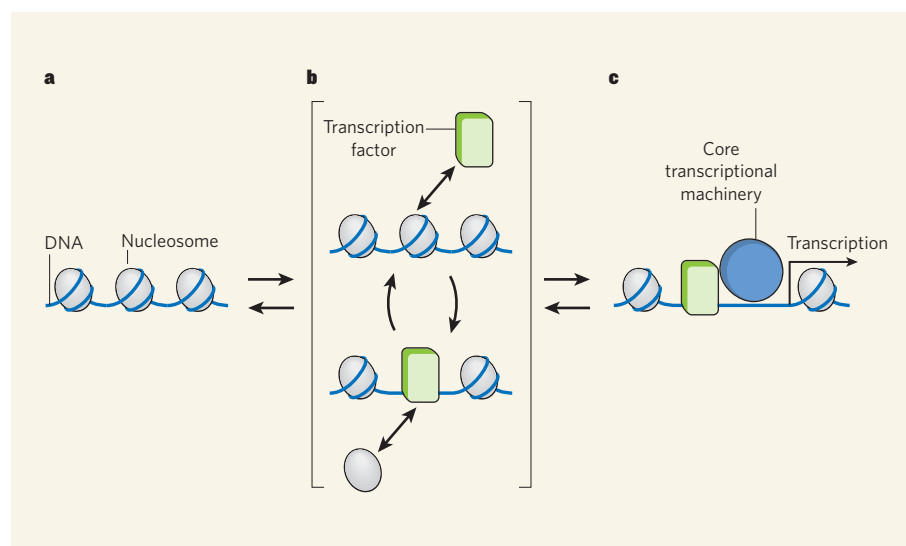


Figure 1 | Well-balanced gene expression. Transcription factors can activate the expression of genes by binding to certain regulatory regions on the genome. Lickwar *et al.*² studied, at the genomic scale, the binding dynamics of one of these proteins, and propose the following model for transcription-factor function. **a**, Most of the genome is wrapped around histone proteins to form nucleosomes, the basic units of DNA packaging into chromatin. **b**, Nucleosomes and transcription factors compete for binding to some regulatory regions. Transcription-factor binding to these regions occurs in short pulses, which are not sufficient for efficient transcription of the gene into RNA. **c**, When the transcription factor binds to its target site for longer periods, it recruits the core transcriptional machinery required to start transcription, leading to high transcription rates.

some genomic sites, it was slowly incorporated into other sites, which suggested that Rap1 binding to the first sites was less stable than to the other sites.

The researchers then applied a mathematical model⁵ to estimate the rate of Rap1 turnover at more than 400 target genes. They found that, among genes with high Rap1 occupancy, those with slower Rap1 turnover showed higher transcription levels than those with faster Rap1 turnover. That is, the transcription level depended on how long Rap1 remained bound. Of note, the genomic sites that exhibited fast Rap1 turnover in this analysis² have previously been reported^{5,6} to have fast turnover rates of nucleosomes (protein complexes around which DNA is packaged) and of the general transcription factor TBP (which facilitates the binding of the core transcriptional machinery). Overall, the results are consistent with those of other studies^{7,8} that showed that transcription factors and nucleosomes compete for some genomic sites and that this competition leads to inefficient transcription.

Lickwar *et al.* suggest a model for the binding dynamics of transcription factors that activate transcription (Fig. 1). In this model, on binding to a target site, the factor has to recruit the core transcriptional machinery. This process takes some time. Therefore, if the factor's binding to the DNA site is unstable, it will not lead to productive transcription. Indeed, it has been shown⁹ that short, repeated pulses of Msn2 — another transcription factor — into the cell nucleus do not activate target genes, whereas longer pulses do. Therefore, for transcription factors to be effective activators, they require stable binding to their target DNA.

Moreover, the researchers speculate that a constant turnover or 'treadmilling' of nucleosomes and transcription factors acts as a distinct mechanism for transcriptional regulation. Unlike static gene repression¹⁰, in which transcription is prevented by the nucleosome's protection of DNA, a site that has a treadmilling transcription factor is poised for activation. When, somehow, the nucleosome is removed or its affinity for DNA is decreased, the factor can quickly achieve stable binding to its target sequence and so activate the gene's transcription. Several mechanisms would allow for the targeted eviction of nucleosomes, including chromatin-remodelling enzymes (which move nucleosomes on DNA), chemical modifications of histones (the protein components of nucleosomes) or the replacement of certain histones with specific variants.

Lickwar and colleagues' study² explains how different regulatory regions can present similar levels of transcription-factor occupancy and different transcriptional levels. But it also raises further exciting questions. Do the different turnover rates of transcription factors play a key part in gene regulation? Or do they just reflect some other aspects of the transcriptional

process, such as stabilization of protein–DNA binding by interactions with the transcriptional machinery? Are nucleosomes the only competition for factor binding to DNA, or is competition with other transcription factors important too?

To fully understand how transcription factors work, we should consider not only their overall binding occupancy, but also their binding dynamics. This line of research will form the basis for a much-needed quantitative understanding of transcription regulation kinetics. ■

Tommy Kaplan is in the Department of Molecular and Cell Biology, California Institute of Quantitative Biosciences, University of California, Berkeley, Berkeley, California 94720, USA, and at the School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem

91904, Israel. **Nir Friedman** is at the School of Computer Science and Engineering, and at the Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.

e-mails: tommy@cs.huji.ac.il; nir@cs.huji.ac.il

1. Macarthur, S. *et al.* *Genome Biol.* **10**, R80 (2009).
2. Lickwar, C. R., Mueller, F., Hanlon, S. E., McNally, J. G. & Lieb, J. D. *Nature* **484**, 251–255 (2012).
3. Lieb, J. D., Liu, X., Botstein, D. & Brown, P. O. *Nature Genet.* **28**, 327–334 (2001).
4. Schermer, U. J., Korber, P. & Hörz, W. *Mol. Cell* **19**, 279–285 (2005).
5. Dion, M. F. *et al.* *Science* **315**, 1405–1408 (2007).
6. van Werven, F., van Teeffelen, H., Holstege, F. & Timmers, H. *Nature Struct. Mol. Biol.* **16**, 1043–1048 (2009).
7. Polach, K. J. & Widom, J. *J. Mol. Biol.* **254**, 130–149 (1995).
8. Voss, T. C. *et al.* *Cell* **146**, 544–554 (2011).
9. Hao, N. & O'Shea, E. K. *Nature Struct. Mol. Biol.* **19**, 31–39 (2011).
10. Hager, G. L., McNally, J. G. & Misteli, T. *Mol. Cell* **35**, 741–753 (2009).

ASTROPHYSICS

Fresh light on stardust

Ageing stars produce elements vital for life and disperse them into space on stellar winds. The discovery of large dust grains in the vicinity of cool giant stars sheds light on the mechanisms that drive such winds. SEE LETTER P.220

SUSANNE HÖFNER

Chemical elements that are crucial for building Earth-like planets and living organisms have their origin in ageing stars and stellar deaths. The nuclear processes that create these elements are well understood, but the mechanisms that transport them to the stellar surfaces and out into interstellar space are still a matter of intensive research. On page 220 of this issue, Norris *et al.*¹ report the detection of silicate particles about 600 nanometres in diameter in the immediate vicinity of several cool giant stars. This result confirms the predictions of models² that explain how gas can escape stellar gravity and become part of the cosmic-matter cycle.

Stellar explosions known as supernovae have considerable input into the production and dispersion of heavy elements (those heavier than helium), but they are not the sole contributors. Stars, including the Sun, release continuous outflows of gas, called stellar winds, for most of their lives. As stars evolve into cool giants and supergiants, stellar winds become increasingly effective in transporting matter out of stellar-gravity wells, enriching the surrounding interstellar medium with newly produced chemical elements. Winds of ageing low- and intermediate-mass stars, such as those observed by Norris *et al.*¹, lead to a runaway mass-loss process that eventually

stops the stars' nuclear processes and turns them into white dwarfs.

The mechanisms that drive gas away from stars differ, depending on a star's surface temperature, mass, luminosity, chemical composition and magnetic field. Winds of cool, luminous giants are presumably triggered by radiative acceleration of dust grains that form in the extended stellar atmospheres. Momentum is transferred from the photons emitted by the star to the dust grains through photon absorption and scattering, and is subsequently redistributed to the more numerous gas particles by collisions with the dust grains (Fig. 1). Because the star's photons are predominantly directed away from the stellar surface, the flow of gas and dust also follows this pattern.

Although the physical principles of dust-driven winds are reasonably well understood, there is currently no consensus on which types of grain are driving the outflows. However, some basic features are known. First, the mass of the gas that is pushed outwards by the dust is more than 100 times higher than the collective mass of the dust particles. This requires grains made from abundant materials that have large radiative cross sections to drive the winds. Second, the grains have to form close to the star to trigger the outflows. This distance is limited by how far shock waves, caused by pulsation and convection in the stellar interiors, can lift gas above the stellar surface, at

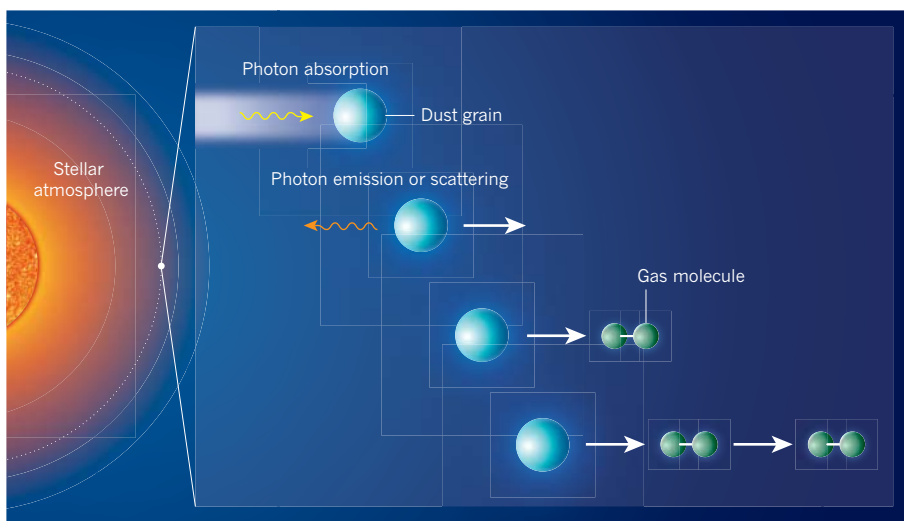


Figure 1 | Dust-driven stellar winds. Dust grains forming in the atmosphere of a cool luminous star are accelerated away (white arrow) from the star through absorption and emission or scattering of stellar photons. By subsequently colliding with molecules in the surrounding gas, the grains accelerate the molecules, make them collide with other gas molecules and trigger an outflow of gas, or stellar wind. Norris and colleagues' study¹ of the immediate vicinity of several cool giant stars provides information on the sizes and material properties of the grains that drive stellar winds.

which temperatures are typically 3,000 kelvin, too high for solid particles to form. The strong flux of stellar photons — which not only accelerates but also heats the grains — severely limits the types of dust material that can form and survive at distances at which winds are triggered.

On the basis of element abundances and distinctive spectral features, magnesium iron silicates of the olivine and pyroxene type have long been considered to be promising candidates for wind drivers³. But detailed models demonstrate⁴ that silicate grains have to be almost devoid of iron to exist at the distance from the stars at which the winds are accelerated, in contrast to earlier assumptions³. Because such particles are highly transparent at near-infrared wavelengths, at which most of the stellar flux is emitted, photon absorption will be insufficient to trigger an outflow. If, however, such iron-free silicate grains grow to 100–1,000 nm in radius, photon scattering dominates over absorption, providing sufficient radiative acceleration for driving winds².

It is therefore of note that silicate grains collected in the Solar System by NASA's Stardust mission⁵ tended to be iron-poor. However, whether this reflects the original composition of grains produced in previous generations of stars before the Solar System formed, or whether stardust is heavily modified in interstellar space and during planetary-system formation, is unclear. To test theoretical predictions, the size and composition of dust grains have to be measured in the immediate vicinities of the stars in which they are formed.

By analysing starlight scattered on dust grains close to several cool giants, Norris

*et al.*¹ found silicate particles with diameters of 600 nm at distances of about two stellar radii, in accordance with model predictions². This result was made possible by a clever combination of advanced instrumentation and observational methods. The authors used measurements of polarization (the direction in which light's electric field oscillates) to identify photons scattered by dust and to determine the size of the grains, and employed interferometric

techniques to obtain high-spatial-resolution images of the close stellar environments in which the winds are triggered.

Identifying wind-driving dust grains in cool giants is a step towards a comprehensive picture of stellar-mass loss, which is essential for understanding ageing stars and their role in the cosmic-matter cycle. Stellar winds strongly affect the evolution of low- and intermediate-mass stars. These stars are probably the progenitors of type Ia supernovae, which are a cornerstone of cosmological studies^{6,7}. Reliable mass-loss rates and dust yields are essential ingredients for understanding the chemical evolution of the Milky Way and other galaxies, and for estimating the effects of interstellar dust on the light of distant galaxies, which is used in studies of the early Universe. The advanced technique used by Norris *et al.*¹ for studying the properties of stardust at its place of origin provides crucial input to these fields of research. ■

Susanne Höfner is in the Department of Physics and Astronomy, Uppsala University, SE-75120 Uppsala, Sweden.
e-mail: hoefner@astro.uu.se

1. Norris, B. R. M. *et al.* *Nature* **484**, 220–222 (2012).
2. Höfner, S. *Astron. Astrophys.* **491**, L1–L4 (2008).
3. Gail, H.-P. & Sedlmayr, E. *Astron. Astrophys.* **347**, 594–616 (1999).
4. Woitke, P. *Astron. Astrophys.* **460**, L9–L12 (2006).
5. Zolensky, M. E. *et al.* *Science* **314**, 1735–1739 (2006).
6. Riess, A. G. *et al.* *Astron. J.* **116**, 1009–1038 (1998).
7. Perlmutter, S. *et al.* *Astrophys. J.* **517**, 565–586 (1999).

MICROBIOLOGY

Adaptation by target remodelling

Bacteria direct their movement in response to certain chemicals by controlling the rotation of whip-like appendages called flagella. The sensitivity of the response can be adjusted at the signal's target, the flagellar motor. [SEE LETTER P.233](#)

GERALD L. HAZELBAUER

When we switch on a lamp in a dark room, our eyes adapt quickly to the lighting. Similarly, in a process known as chemotaxis, specific sensory systems in bacteria detect and adapt to changes in the concentration of molecules such as nutrients, so that the microbes can control the motion of their flagella and swim to more favourable environments. Many adaptive mechanisms act on cell-surface receptor proteins that detect stimuli and generate signals across the cell membrane and within the cytoplasm. Indeed,

the sensitivity of the bacterial chemotaxis machinery is regulated by addition or removal of methyl groups on the chemical-sensing receptors (chemoreceptors)¹. But, as Yuan *et al.*² report on page 233 of this issue, the machinery can also adapt by remodelling the target of signalling — the flagellar motor. Motor adaptation differs from receptor adaptation in its time frame, mechanism and function.

In the bacterial chemotaxis system, different chemoreceptors can bind to either attractant or repellent ligand molecules, and by doing so they alter the activity of a receptor-associated

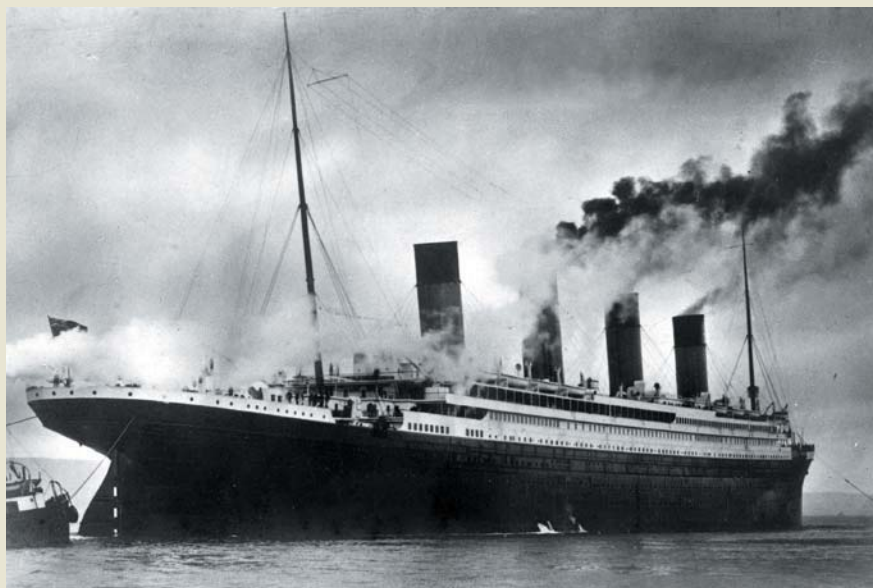


100 Years Ago

The loss of the “Titanic”

The terrible loss of life on account of the disaster to the *Titanic* has directed emphatic attention to various aspects of the employment of wireless telegraphy in times of crisis at sea. The point which is at the moment attracting most of the public attention is that of the erroneous messages, or alleged messages, which appeared in the newspapers in the day or two following the disaster ... All this raises more prominently than ever the chaotic condition of wireless telegraphy in the United States ... [T]he most urgent call for help will pass unheeded if none of the operators on the ships within hail are on duty. In fact, it seems to have been a mere chance that the *Carpathia* operator was at his apparatus at the time the *Titanic* called. On ships that carry only one operator — and very few carry more — the man cannot always be on the look-out ...

Engineering aspects of the disaster are discussed in the leading article in *Engineering* for April 19 ... [S]everal questions present themselves as ripe for discussion and settlement. The effect of centre-line or longitudinal wing bulkheads is one of these. Such have advantages in confining any water admitted to a part of the width, but have disadvantages even from the point of view of stability under disastrous conditions. The effect of impact on the superstructure of very large ships will have to be considered. In such ships it has become a practice to have two or three decks above the moulded structure. Would inertia have effects somewhat similar to those experienced in railway collisions, in which the body of the carriage is driven from the



Titanic sank on 15 April 1912 on its maiden voyage, after hitting an iceberg off Newfoundland.

under-frame? ... The engineers of the ship have all been lost — their claim to recognition is the simplest and best; they did their duty to the end.

From *Nature* 25 April 1912

At the time when the *Titanic* was lost the standing Advisory Committee appointed by the Board of Trade under the provisions of Merchant Shipping Acts was engaged in the reconsideration of the regulations for boats and life-saving appliances ... The main recommendations of [their] report may be summarised. First, it is recognised that “the stability and seaworthy qualities of the vessel itself” must be regarded as of primary importance. This includes the question of watertight subdivision, now under investigation by a special committee. Second, as regards boats and life-saving appliances it is recommended that accommodation should be provided for the *total number* of persons which each *foreign-going* passenger steamship is licensed to carry ... One of [the committee’s] most valuable recommendations is that proposing to extend the present regulations and to prescribe to those in charge of ships the necessity for

proceeding at moderate speed “at night in the known vicinity of ice.”

From *Nature* 29 August 1912

It behoves surely men of science to ask the question whether we have not reached the imperative limits of that false security which the “practical man” is wont to feel in his contempt for scientific “theory”; and further, whether the time has not therefore come for legislation requiring commanders of the largest ocean-going steamers to hold a diploma, guaranteeing such a systematic course of study (say in a class at Greenwich or Kensington) in marine physiography and the elementary laws of mechanics as would quicken their imagination as to the uncertainty and the magnitude of the risks to be run in an abnormally ice-drifted sea. Lord Mersey’s report may whitewash the facts, but the facts *en évidence* remain; and the chain of cause and effect in the lamentable and tragic loss of the *Titanic* leads us in the last resort to the notorious contempt for scientific acquaintance with the facts and laws of nature on the part of the “practical man”.

From *Nature* 12 September 1912

kinase enzyme² (Fig. 1). The kinase adds a phosphate group to the protein CheY, which can bind only in its phosphorylated form (CheY-P) to FliM, a protein component of the flagellar rotary motor. Binding leads to a switch in flagellar rotation from anticlockwise — the default state — to clockwise. As anticlockwise rotation powers forward swimming (‘runs’) and clockwise rotation causes abrupt directional changes (‘tumbles’), their alternation generates a three-dimensional random walk. When the bacterium swims towards increasing

concentrations of attractant, the interaction of attractant with its receptors reduces kinase activity, and thus CheY-P levels and the probability of tumbles, biasing the random walk in favourable directions.

But when ligand concentration remains constant, the system adapts by reverting to its null state, so that it can then respond to any additional concentration changes. Adaptation occurs through the addition or removal of methyl groups to the receptors in response to occupancy by attractants or repellents.

Methylation and demethylation shift receptor conformation and thus kinase activity, CheY-P levels, rotational bias and swimming behaviour in the opposite direction to attractant and repellent occupancy, respectively. These receptor adaptive changes occur more slowly than those generated by ligand binding, and this difference provides the bacterium with a means of comparing current and recent ligand concentrations through the opposing influences of receptor occupancy and the extent of methylation. The overall result is

that the microbe can sense temporal changes in the concentration of relevant molecules and therefore can swim in the 'right' direction.

The flagellar motor is extremely sensitive to small variations in CheY-P concentration³, and this poses a challenge to the system because the steady-state CheY-P concentration varies among individual cells⁴. Given this variation, some cells would be expected to run only or to tumble only, and therefore to be unable to perform proper chemotaxis. As this does not seem to occur, a mechanism must exist to keep a correspondence between steady-state CheY-P levels and the narrow CheY-P concentration range over which the rotational switch is sensitive. However, previous research⁵ did not find one possible mechanism — feedback from the flagellar rotary motor to the kinase.

Yuan *et al.*² investigated an alternative possibility: could the motor itself adapt to changes in CheY-P concentration? The authors used a strain of *Escherichia coli* that lacked the enzymes that carry out receptor methylation and demethylation, as otherwise the process under study would have been masked by the rapid and complete adaptation mediated by these enzymes. When the bacteria were stimulated by an attractant, the authors observed that the CheY-P concentration and the motor's clockwise bias both decreased rapidly, as expected. However, over the next few minutes, even though CheY-P levels remained unchanged, the clockwise bias gradually increased to an intermediate steady-state value that was lower than in the absence of attractant. A similar slow, partial adaptation of the flagellar motor was reported^{6,7} 25 years ago in bacteria lacking the receptor methylation and demethylation enzymes.

The researchers² also measured the clockwise bias of flagellar motors in individual bacterial cells before stimulation, at maximum response and after partial adaptation. In the presence of persistently decreased CheY-P concentration, they documented a shift in the motor's sensitivity to CheY-P to a lower, but still narrow, concentration range. When they analysed these data using a mathematical model for cooperative motor switching, the results indicated that the adaptation probably reflected an increased number of binding sites for CheY-P in the flagellar motor — in other words, more copies of FliM.

To test this hypothesis, the authors expressed a modified version of FliM (fused to a yellow fluorescent protein) in the *E. coli* cells, and then observed these cells using total internal reflection fluorescence (TIRF) microscopy. The researchers did indeed detect increases of up to about 25% in the number of FliM copies (from 34 to 42 copies) in flagellar motors upon partial adaptation. This result is consistent with a previous finding⁸ of CheY-P-dependent turnover of FliM units, and suggests a function for such a process: the addition of FliM copies to the flagellar motor would increase the

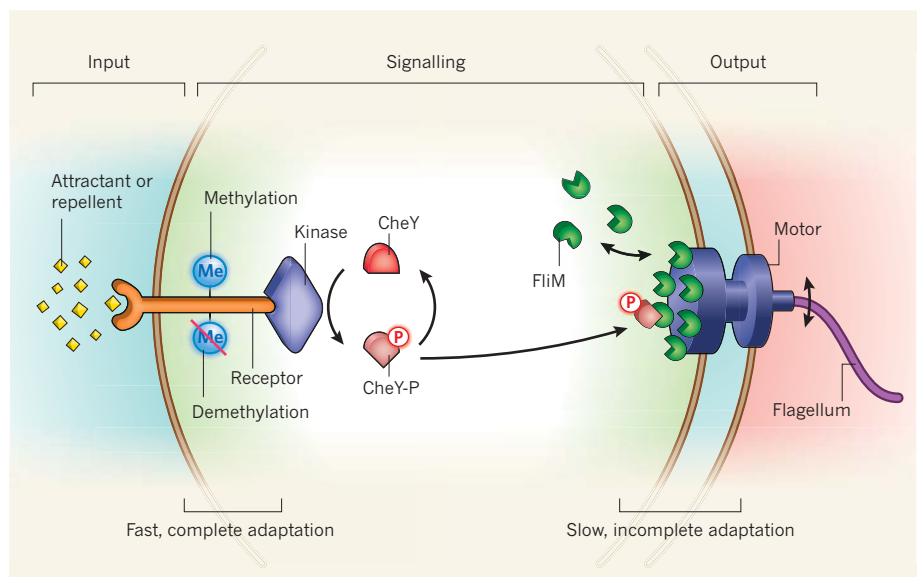


Figure 1 | Microbial sense and sensibility. Protein receptors on a bacterial cell's surface bind to specific attractant or repellent molecules, which act as inputs for the bacterium's sensory system. In the absence of attractant, a receptor-associated kinase enzyme donates a phosphoryl group (P) to the protein CheY. Phosphorylated CheY (CheY-P) is an intracellular signal that interacts with the protein FliM at the flagellar motor. This interaction generates a change in rotational direction of the motor — the system's output. Upon attractant binding, the receptor undergoes a conformational change that reduces kinase activity and consequently the levels of CheY-P and the probability of a change in the direction of rotation. Binding of repellent has the opposite effects. If attractant or repellent concentrations do not change for a while, the sensory system 'resets' rapidly and completely, within seconds, to a null state through receptor methylation (for attractant) or demethylation (for repellent) at the input end. Yuan *et al.*² find that the system also slowly and partially adapts at the output end to different, persistent levels of CheY-P. It does so by changing the number of FliM copies in the flagellar motor, thus shifting the narrow sensitivity range of the motor to correspond to different steady-state levels of CheY-P.

probability of CheY-P binding to the motor and therefore the sensitivity of the rotational switch.

Yuan and colleagues' results² generate questions for future research. For instance, how does the flagellar motor sense changes in the level of CheY-P? And how does sensing lead to an increased number of FliM copies? How few and how many copies can be accommodated, and what structural interactions between FliM units allow such variability? Answering these questions will be challenging, but insight could be gained by identifying and characterizing mutant bacteria that are defective in motor adaptation, and by extending high-resolution fluorescence analyses such as the TIRF microscopy experiments done by the authors. Full answers are likely to require both biochemical and structural studies.

It also needs emphasizing that motor adaptation is a slow process, occurring over minutes. So, in the absence of methylation-dependent adaptation of the receptors, which occurs within seconds, bacteria cannot perform chemotaxis^{9,10}. Nonetheless, motor adaptation is likely to be crucial for accommodating stochastic variations in the dosages of protein components of the chemotaxis sensory machinery, such as CheY. The adjustment of the sensitivity of a signalling system by adaptive remodelling of its output end is a

tantalizing observation, because experience tells us that if a new strategy is discovered in one biological system, then there are undiscovered examples in other systems. Therefore, as the authors suggest, other biological molecular machines may adapt to changes in signal levels by resetting the sensitivity of their response. ■

Gerald L. Hazelbauer is in the Department of Biochemistry, School of Medicine and the College of Agriculture, Foods and Natural Resources, University of Missouri, Columbia, Missouri 65211, USA.
e-mail: hazelbauer@missouri.edu

1. Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. *Trends Biochem. Sci.* **33**, 9–19 (2008).
2. Yuan, J., Branch, R. W., Hosu, B. G. & Berg, H. C. *Nature* **484**, 233–236 (2012).
3. Cluzel, P., Surette, M. & Leibler, S. *Science* **287**, 1652–1655 (2000).
4. Alon, U., Surette, M. G., Barkai, N. & Leibler, S. *Nature* **397**, 168–171 (1999).
5. Shimizu, T. S., Delalez, N., Pichler, K. & Berg, H. C. *Proc. Natl Acad. Sci. USA* **103**, 2093–2097 (2006).
6. Stock, J., Kersulis, G. & Koshland, D. E. Jr *Cell* **42**, 683–690 (1985).
7. Segall, J. E., Block, S. M. & Berg, H. C. *Proc. Natl Acad. Sci. USA* **83**, 8987–8991 (1986).
8. Delalez, N. J. *et al. Proc. Natl Acad. Sci. USA* **107**, 11347–11351 (2010).
9. Hazelbauer, G. L., Park, C. & Nowlin, D. M. *Proc. Natl Acad. Sci. USA* **86**, 1448–1452 (1989).
10. Wolfe, A. J. & Berg, H. C. *Proc. Natl Acad. Sci. USA* **86**, 6973–6977 (1989).

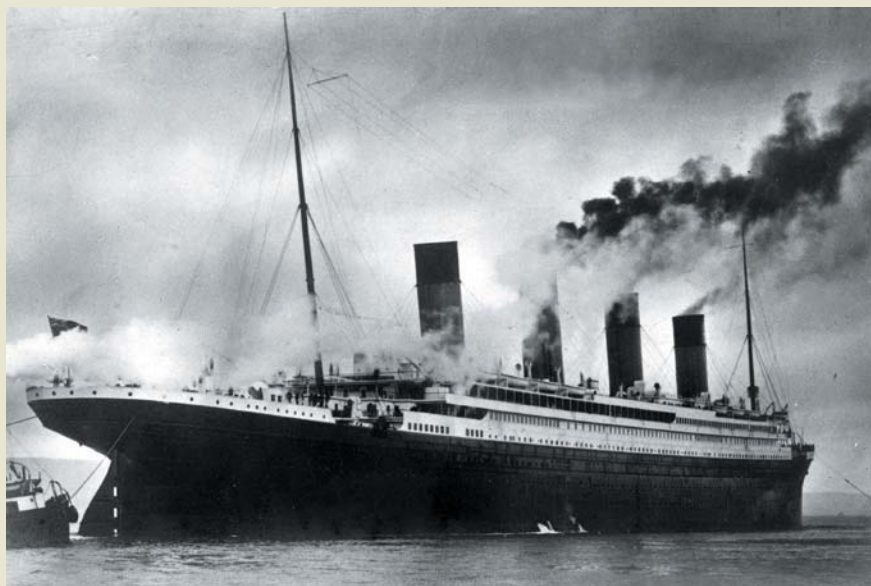


100 Years Ago

The loss of the “Titanic”

The terrible loss of life on account of the disaster to the Titanic has directed emphatic attention to various aspects of the employment of wireless telegraphy in times of crisis at sea. The point which is at the moment attracting most of the public attention is that of the erroneous messages, or alleged messages, which appeared in the newspapers in the day or two following the disaster ... All this raises more prominently than ever the chaotic condition of wireless telegraphy in the United States ... [T]he most urgent call for help will pass unheeded if none of the operators on the ships within hail are on duty. In fact, it seems to have been a mere chance that the Carpathia operator was at his apparatus at the time the Titanic called. On ships that carry only one operator — and very few carry more — the man cannot always be on the look-out ...

Engineering aspects of the disaster are discussed in the leading article in *Engineering* for April 19 ... [S]everal questions present themselves as ripe for discussion and settlement. The effect of centre-line or longitudinal wing bulkheads is one of these. Such have advantages in confining any water admitted to a part of the width, but have disadvantages even from the point of view of stability under disastrous conditions. The effect of impact on the superstructure of very large ships will have to be considered. In such ships it has become a practice to have two or three decks above the moulded structure. Would inertia have effects somewhat similar to those experienced in railway collisions, in which the body of the carriage is driven from the



Titanic sank on 15 April 1912 on its maiden voyage, after hitting an iceberg off Newfoundland.

under-frame? ... The engineers of the ship have all been lost — their claim to recognition is the simplest and best; they did their duty to the end.

From *Nature* 25 April 1912

At the time when the Titanic was lost the standing Advisory Committee appointed by the Board of Trade under the provisions of Merchant Shipping Acts was engaged in the reconsideration of the regulations for boats and life-saving appliances ... The main recommendations of [their] report may be summarised. First, it is recognised that “the stability and seaworthy qualities of the vessel itself” must be regarded as of primary importance. This includes the question of watertight subdivision, now under investigation by a special committee. Second, as regards boats and life-saving appliances it is recommended that accommodation should be provided for the total number of persons which each foreign-going passenger steamship is licensed to carry ... One of [the committee’s] most valuable recommendations is that proposing to extend the present regulations and to prescribe to those in charge of ships the necessity for

proceeding at moderate speed “at night in the known vicinity of ice.”

From *Nature* 29 August 1912

It behoves surely men of science to ask the question whether we have not reached the imperative limits of that false security which the “practical man” is wont to feel in his contempt for scientific “theory”; and further, whether the time has not therefore come for legislation requiring commanders of the largest ocean-going steamers to hold a diploma, guaranteeing such a systematic course of study (say in a class at Greenwich or Kensington) in marine physiography and the elementary laws of mechanics as would quicken their imagination as to the uncertainty and the magnitude of the risks to be run in an abnormally ice-drifted sea. Lord Mersey’s report may whitewash the facts, but the facts en evidence remain; and the chain of cause and effect in the lamentable and tragic loss of the Titanic leads us in the last resort to the notorious contempt for scientific acquaintance with the facts and laws of nature on the part of the “practical man”.

From *Nature* 12 September 1912

kinase enzyme² (Fig. 1). The kinase adds a phosphate group to the protein CheY, which can bind only in its phosphorylated form (CheY-P) to FliM, a protein component of the flagellar rotary motor. Binding leads to a switch in flagellar rotation from anticlockwise — the default state — to clockwise. As anticlockwise rotation powers forward swimming (‘runs’) and clockwise rotation causes abrupt directional changes (‘tumbles’), their alternation generates a three-dimensional random walk. When the bacterium swims towards increasing

concentrations of attractant, the interaction of attractant with its receptors reduces kinase activity, and thus CheY-P levels and the probability of tumbles, biasing the random walk in favourable directions.

But when ligand concentration remains constant, the system adapts by reverting to its null state, so that it can then respond to any additional concentration changes. Adaptation occurs through the addition or removal of methyl groups to the receptors in response to occupancy by attractants or repellents.

Methylation and demethylation shift receptor conformation and thus kinase activity, CheY-P levels, rotational bias and swimming behaviour in the opposite direction to attractant and repellent occupancy, respectively. These receptor adaptive changes occur more slowly than those generated by ligand binding, and this difference provides the bacterium with a means of comparing current and recent ligand concentrations through the opposing influences of receptor occupancy and the extent of methylation. The overall result is

High-valent organometallic copper and palladium in catalysis

Amanda J. Hickman¹ & Melanie S. Sanford¹

Copper and palladium catalysts are critically important in numerous commercial chemical processes. Improvements in the activity, selectivity and scope of these catalysts could drastically reduce the environmental impact, and increase the sustainability, of chemical reactions. One rapidly developing strategy for achieving these goals is to use 'high-valent' organometallic copper and palladium intermediates in catalysis. Here we describe recent advances involving both the fundamental chemistry and the applications of these high-valent metal complexes in numerous synthetically useful catalytic transformations.

Homogeneous copper- and palladium-catalysed reactions are widely used for the construction of important organic molecules, including pharmaceuticals^{1–3}, commodity chemicals⁴ and polymers⁵. The development of copper and palladium catalysis has been inextricably linked because both metals have been used extensively in the construction of similar types of carbon–carbon and carbon–heteroatom bonds. Furthermore, advancements in, and insights into, copper chemistry have often spurred improvements in palladium-catalysed processes, and vice versa, leading to a wide range of robust, synthetically valuable and often complementary synthetic methods. This Review focuses particularly on an area that has progressed tremendously over the past decade: the use of high-valent copper and palladium complexes in catalysis (see Box 1 for definitions of 'high-valent' copper and palladium).

History and importance of copper catalysis

Copper is an inexpensive, earth-abundant, non-toxic metal that has found widespread application in homogeneous catalysis. For example, copper-catalysed cross-coupling reactions have been extensively studied since their discovery at the turn of the twentieth century^{6,7}. Such reactions serve as versatile methods for synthesizing biaryl linkages as well as for constructing the carbon–heteroatom bonds of aryl amines, aryl ethers and aryl thioether derivatives. The alkylation of carbon electrophiles with organometallic copper compounds (organocuprates) is another classic reaction in organic synthesis³. Diverse carbon–carbon bond-forming reactions of organocuprates have been developed over the past 70 years, and these transformations are featured in most introductory undergraduate textbooks on organic chemistry.

High-valent copper intermediates (that is, organometallic copper(III) (Cu(III)) species) have long been proposed to have a role in both copper-catalysed cross-coupling and organocuprate reactions¹⁰. In particular, carbon–carbon and/or carbon–heteroatom bond formation from an organo-Cu(III) species has been invoked as the product-releasing step of many of these transformations^{8,9,11,12}. However, the proposed high-valent copper intermediates have eluded detection for decades, and as a result there has been considerable controversy over the mechanistic details of both organocuprate additions and copper-catalysed cross-coupling¹¹. Indeed, until very recently, these two transformations were among the least mechanistically understood synthetic methods in organometallic chemistry. These mechanistic questions and controversy have provided tremendous motivation for probing the accessibility and reactivity of organo-Cu(III) species. A deeper mechanistic understanding

of their chemistry promises to allow the development of improved copper catalysts for known reactions as well as to inspire novel copper-catalysed transformations. In the section on high-valent copper below, we discuss recent progress on many of the vital questions in this area (discussed in detail in Box 1).

History and importance of palladium catalysis

Although copper-mediated cross-coupling methods were the first of their kind, today cross-coupling has become synonymous with a different metal: palladium. Well-defined palladium-catalysed cross-coupling reactions were first developed in the 1970s, and they quickly surpassed copper-based methods in both popularity and scope. These reactions have transformed the way organic chemists approach the construction of bonds in complex molecules^{1,5,13}, and the wide-ranging impact of this methodology was recognized in the awarding of the Nobel Prize in Chemistry in 2010.

The rapid success of palladium-catalysed cross-coupling is due, in large part, to extensive and systematic investigations of reaction mechanisms. Mechanistic analysis has revealed that nearly all of these processes involve catalysis by 'low-valent' palladium (that is, palladium in the 0 (Pd(0)) or +2 (Pd(II)) oxidation states). For many reactions, Pd(0) and Pd(II) catalytic intermediates have been well characterized and the steric and electronic influence of supporting ligands on catalysis has been studied in detail¹⁴. Such mechanistic studies have been crucial in the development of new catalyst structures and novel transformations with wide scope and mild reaction conditions^{1,13,14}.

Although catalysis by low-valent palladium is ubiquitous and extremely synthetically useful, it has several important limitations that stem from the fundamental properties of organo-Pd(II) complexes. These include limited reactivity to forming certain important types of chemical bond (for example carbon–halogen and carbon–CF₃ linkages) and a high susceptibility to decomposition pathways such as β -hydride elimination. These challenges have motivated the study of catalysis by high-valent palladium as a potentially complementary mechanistic pathway. The first 30 years of palladium chemistry were dominated by the use of low-valent palladium, but over the past decade the unique reactivity of Pd(III) and Pd(IV) intermediates has increasingly been recognized and exploited in catalysis. In the section on high-valent palladium below, we present recent advances in the field that demonstrate the relevance and utility of high-valent palladium in diverse catalytic applications (discussed in detail in Box 1).

¹University of Michigan, Department of Chemistry, 930 North University Avenue, Ann Arbor, Michigan 48109, USA.

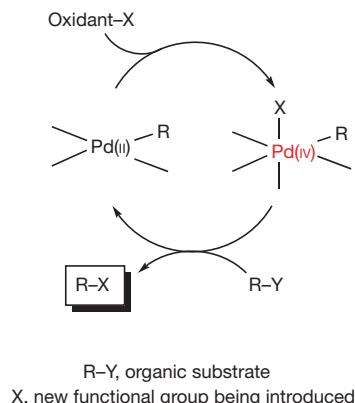
BOX 1

High-valent copper and palladium

In the context of palladium and copper, high-valent compounds include palladium in the +3 or +4 oxidation state (Pd(III) and Pd(IV), respectively) and copper in the +3 oxidation state (Cu(III)). Organometallic complexes of these high-valent metals are defined as molecules that contain copper–carbon or palladium–carbon bonds. Catalysis by high-valent copper or palladium is defined as a catalytic reaction in which the metal is oxidized to form a high-valent organometallic intermediate during the catalytic cycle.

Advances in the chemistry of high-valent copper have enabled researchers to address a number of critical issues, including the synthetic accessibility of organo-Cu(III) complexes, the viability of carbon–carbon and carbon–heteroatom bond formation from discrete organo-Cu(III) species, the catalytic relevance of Cu(III) complexes and the ability to exploit high-valent copper intermediates to improve catalytic reactions and/or discover new reactivity modes.

There has been a renaissance in the chemistry of high-valent palladium over the past decade that has provided key insights into the following issues: the synthetic accessibility of high-valent Pd(III) or Pd(IV) organometallic complexes, the ability of these species to participate in stoichiometric carbon–carbon and carbon–heteroatom bond-forming reactions, the catalytic relevance of these high-valent palladium species and the advantages of high-valent palladium catalysis (in terms of enhanced substrate range, milder reaction conditions and improved chemoselectivity, regioselectivity and/or site selectivity) over more-common low-valent analogues. A representative high-valent palladium catalytic cycle is shown in the figure.



High-valent copper

Before 2000, high-valent organometallic copper complexes were rare. The occasional examples reported in the literature were stabilized by rigid, chelating and/or perfluorinated ligands^{15–17}, as exemplified by compounds 1–3 in Fig. 1a. Although complexes 1–3 are structurally interesting, they do not have the characteristic reactivity that has been attributed to Cu(III) in catalysis. Specifically, they are all inert to carbon–carbon and carbon–heteroatom bond-forming reactions. As such, these compounds were largely considered curiosities, whose relevance to copper catalysis was tenuous at best. The past ten years have seen tremendous developments in this area with the observation and detailed investigation of catalytically relevant organo-Cu(III) species in both carbon–carbon and carbon–heteroatom bond formation. In this section, we will specifically focus on two representative areas: high-valent copper intermediates in carbon–carbon bond-forming reactions of organocuprates and high-valent copper intermediates in carbon–nitrogen and carbon–oxygen couplings.

An early advance in the chemistry of high-valent copper came from investigations of carbon–carbon bond-forming reactions of organocuprates with enones, alkyl halides and allylic electrophiles. Computational studies

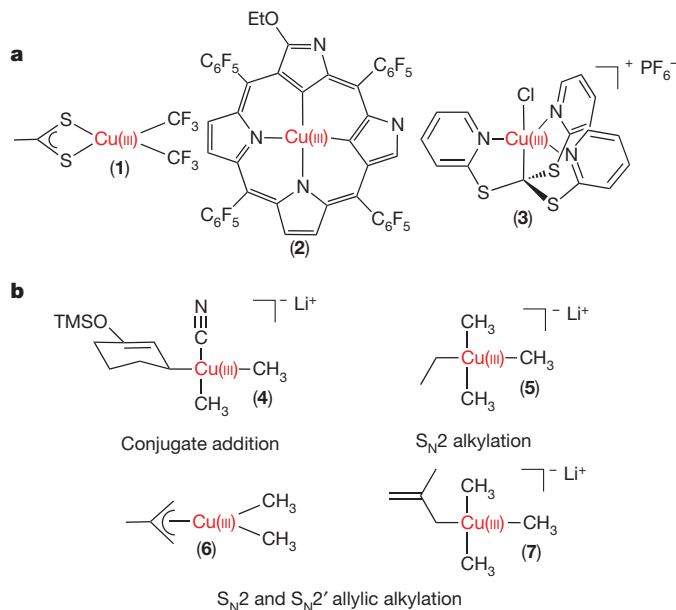


Figure 1 | High-valent copper complexes. **a**, Early examples of isolatable organometallic Cu(III) complexes. **b**, Cu(III) intermediates of organocuprate reactions detected at $-100\text{ }^{\circ}\text{C}$ using rapid-injection NMR spectroscopy. Et, ethyl; TMS, trimethylsilyl.

of these transformations implicated a Cu(III) intermediate in the key carbon–carbon bond-forming step^{18–20}. However, for many years little experimental evidence was available to support this hypothesis, as the putative Cu(III) compounds proved too transient for detection using standard spectroscopic techniques²¹. In 2007, rapid-injection NMR spectroscopy (RI-NMR) was introduced as a method to directly observe Cu(III) species such as 4–7 in real time (Fig. 1b). Remarkably, when generated using RI-NMR, the Cu(III) adducts could be detected and fully characterized at $-100\text{ }^{\circ}\text{C}$ (refs 22–26). Furthermore, on being warmed, these discrete organo-Cu(III) intermediates underwent carbon–carbon bond-forming reductive elimination²².

This field is still in its infancy, but the ability to study directly the reactivity of high-valent organo-Cu(III) species has begun to provide mechanistic insights of direct relevance to copper catalysis. For example, Lewis basic additives such as cyanide, phosphines, pyridines and amines have been known for decades to improve the yield and/or rate of organocuprate conjugate addition reactions^{27,28}. Some researchers have proposed that the primary role of these additives is to enhance the solubility of copper starting materials or intermediates²⁸. By contrast, other groups have speculated that these Lewis bases have a more intimate role in the reaction mechanism by binding to copper and tuning its reactivity towards oxidative addition and/or carbon–carbon bond-forming reductive elimination^{19,27}.

RI-NMR has provided a way of interrogating these possibilities directly. A series of Cu(III) complexes of general structure $(\text{CH}_3\text{CH}_2)(\text{CH}_3)_2\text{Cu(III)LB}$, where LB denotes a Lewis basic ligand, were prepared using this technique and evaluated as models for conjugate addition intermediates²⁹. The nature of LB was found to have a profound influence on the stability of these species. For instance, when LB was pyridine, the Cu(III) complex **8** was a short-lived intermediate at $-100\text{ }^{\circ}\text{C}$ (0.5 h to maximum concentration) (Fig. 2). At this temperature, **8** underwent facile ligand exchange to form the Cu(III) complex $(\text{CH}_3\text{CH}_2)(\text{CH}_3)_3\text{Cu(III)Li}$, as well as carbon–carbon bond-forming reductive elimination to release propane and $(\text{CH}_3)_3\text{Cu(I)}_2\text{Li}$. By contrast, under otherwise analogous conditions the dimethylaminopyridine complex **9** was very stable at $-100\text{ }^{\circ}\text{C}$ (Fig. 2). Minimal (<10%) reductive elimination was detected, and ligand exchange to form $(\text{CH}_3\text{CH}_2)(\text{CH}_3)_3\text{Cu(III)Li}$ was not observed in this system.

Reference 29 clearly shows that Lewis basic ligands drastically influence the relative and absolute rates of carbon–carbon coupling at Cu(III)

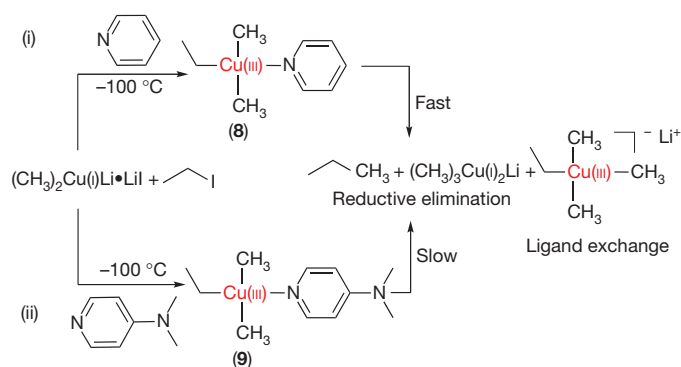


Figure 2 | RI-NMR studies of the effect of Lewis bases on the reactivity of Cu(III) complexes 8 and 9. Path (i): pyridine-containing intermediate 8 is short lived and undergoes ligand exchange to form $(\text{CH}_3\text{CH}_2)(\text{CH}_3)_3\text{Cu(III)Li}$ as well as carbon–carbon bond-forming reductive elimination. Path (ii): by contrast, dimethylaminopyridine-containing intermediate 9 is stable at -100°C under analogous conditions.

centres. In the future, more-quantitative RI-NMR analysis of reaction kinetics and ligand electronic/steric effects should provide further insights about the rate- and selectivity-determining steps of conjugate addition and other organocuprate reactions. Such studies will also be invaluable in establishing the role of Cu(III) in the reactions of organocuprates with other electrophiles (for example acyl halides, carbonyl compounds and cyclopropanes). Furthermore, they will provide a mechanistic platform for rationally designing new synthetic methods in this area.

The ability to use RI-NMR to observe and study organo-Cu(III) intermediates also has important implications for emerging areas of copper catalysis. For example, the copper-catalysed carbon–hydrogen arylation of anilides with diphenyliodonium trifluoromethanesulphonate, $[\text{Ph}_2\text{I}]\text{OTf}$, was recently reported³⁰. In these reactions, carbon–carbon bond formation occurs *meta* to the amide substituent, site selectivity that is complementary to palladium-catalysed, ligand-directed carbon–hydrogen arylation methods³¹. The mechanism of this reaction remains controversial. Calculations using density functional theory have implicated the intermediacy of a Cu(III)–phenyl complex³²; however, the accessibility of such intermediates has yet to be confirmed experimentally. Alternative mechanisms such as Lewis acid catalysis are also plausible. RI-NMR would serve as a powerful technique for detecting Cu(III) (if present) during catalysis and/or for interrogating stoichiometric reactions of Cu(I) model complexes with $[\text{Ph}_2\text{I}]\text{OTf}$. Such studies could help to clarify the mechanism of this novel transformation as well as to probe the origin of the *meta* selectivity.

A second key advance in catalysis by high-valent copper came in the study of carbon–heteroatom bond formation from Cu(III) intermediates. As representative examples, we highlight recent investigations of the copper-catalysed amination of aryl bromides and of copper-catalysed carbon–hydrogen bond amination and oxygenation. Copper catalysts are well known to promote the amination of aryl boronic acids^{33,34}, aryl halides^{35,36} and carbon–hydrogen bonds^{37–42}. Common nitrogen heterocycles such as pyrazole, pyridone and phthalimide are particularly effective coupling partners, and the reactions often proceed under mild conditions. Many researchers have proposed the intermediacy of Cu(III) in these transformations^{18,43,44}. However, this hypothesis has been the subject of significant debate, and others have argued that single-electron transfer, halide atom transfer or σ -bond metathesis mechanisms at low-valent Cu(I) or Cu(II) are more likely^{45,46}. Until very recently, no Cu(III) catalytic intermediates had been detected, and carbon–nitrogen bond formation from a Cu(III) complex had not been observed directly.

The synthesis of the first isolatable Cu(III)–monoaryl species (10; Fig. 3a), in 2002, was a turning point for this field⁴⁷. Like many of the early examples of organometallic Cu(III) compounds (for example 2 and 3 in Fig. 1), complex 10 is stabilized by an electron-donating macrocyclic ligand. However, unlike its predecessors 10 is remarkably reactive to

carbon–nitrogen bond formation. For example, 10 reacts stoichiometrically with amines such as pyridone, oxazolidinone and acetanilide to form aminated products⁴³ (11; Fig. 3a). Further study showed that less basic amines reacted faster, implicating deprotonation of the amine at or before the rate-determining step of this sequence. Notably, in the absence of amine, 13 was shown to undergo carbon–bromine bond-forming reductive elimination to release aryl bromide 12 when trifluoromethanesulphonic acid is present (Fig. 3b).

The next key question was whether Cu(III) complex 10 and analogues thereof are relevant to catalytic carbon–nitrogen coupling reactions. The copper-catalysed amination of aryl bromide 12 with pyridone was studied to test this possibility³⁵. Remarkably, it was possible to detect a steady-state concentration of Cu(III) intermediate 13 during catalysis by means of ultraviolet–visible and NMR spectroscopy (Fig. 3b). The observation of 13 is consistent with its participation in the turnover-limiting step of the catalytic reaction. Although studies of this one system do not resolve the controversy surrounding the mechanism of all copper-catalysed cross-coupling reactions, they provide the first definitive demonstration that Cu(III) can be catalytically relevant in these transformations.

Pioneering studies have also recently established a role for high-valent copper in certain copper-catalysed carbon–hydrogen bond-functionalization reactions^{37,48}. Before these studies, there was considerable mechanistic uncertainty about these transformations^{38–42,49–53}. Radical pathways initiated by single-electron transfer from amine, enolate and electron-rich arene substrates have frequently been proposed^{42,52}. However, a growing number of examples have been reported with substrates (for example alkynes^{50,51} and electron-deficient arenes^{40,42,52,53}) that are unlikely to participate in such a mechanism. Very recently, the catalytic aerobic carbon–hydrogen oxygenation of macrocycle 14 with methanol (Fig. 3c) was demonstrated³⁷. *In situ* ultraviolet–visible spectroscopic studies revealed the build-up and subsequent decay of Cu(III) complex 13 during the catalytic reaction, implicating this species as a catalytically relevant intermediate. Further kinetic studies suggested that the rates of Cu(II)-mediated carbon–hydrogen cleavage and of carbon–oxygen bond formation from Cu(III) are closely matched, which would explain the observed concentration profile of intermediate 13 during catalysis³⁷.

In summary, ten years ago little was known about the stability and reactivity of high-valent copper complexes. Since then there has been considerable progress, including the first observation and study of carbon–carbon and carbon–heteroatom bond formation from discrete organo-Cu(III) species in stoichiometric and catalytic transformations. Fundamental studies of organo-Cu(III) are beginning to provide greater understanding of mechanism, which in turn should allow the rational development of new synthetic methods⁵⁴.

High-valent palladium

Sporadic reports over the past 50 years have proposed the intermediacy of Pd(IV) in catalysis^{55–57}. However, these proposals were frequently viewed with scepticism owing to a lack of evidence supporting the viability of such species. Thus, a first key challenge was to determine whether it was possible to form, detect and isolate Pd(III) and/or Pd(IV) complexes from the reactions of Pd(II) precursors with oxidants. Early work established the viability of this approach and demonstrated that electron-donating, rigid, multidentate supporting ligands can be used to stabilize high-valent palladium products. For example, in 1988 organometallic Pd(IV) complex 16 (Fig. 4) was prepared through the reaction of 15 (containing the rigid, bidentate 2,2'-bipyridine ligand) with CH_3I (ref. 58). Similarly, organo-Pd(III) dimer 18 was produced by reacting 17 (containing bidentate, electron-donating cyclometalated phosphines) with PhICl_2 (ref. 59). These seminal discoveries have inspired extensive efforts to exploit related intermediates in catalysis. In this section, we specifically focus on progress in two areas: high-valent palladium intermediates in carbon–halogen bond formation and high-valent palladium intermediates in trifluoromethylation reactions.

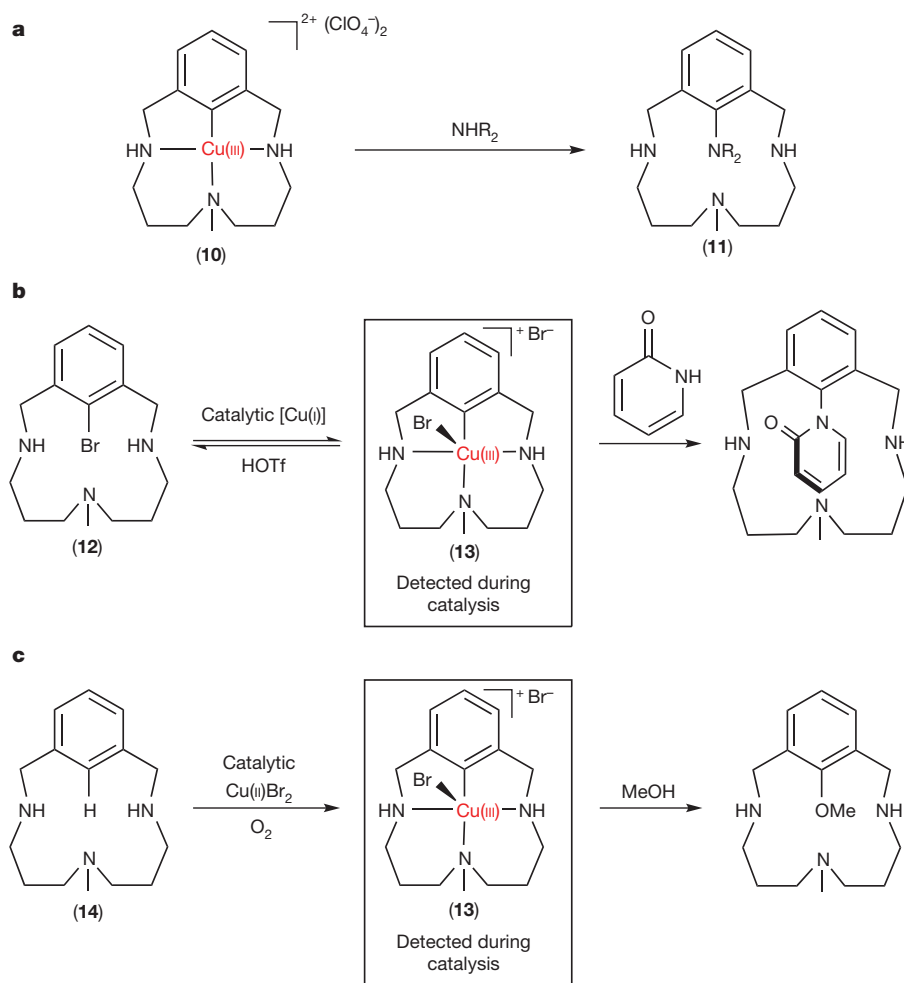


Figure 3 | High-valent copper complexes involved in carbon-heteroatom bond formation. **a**, Stoichiometric carbon–nitrogen bond formation from an isolated organo-Cu(III). NHR₂ represents pyridone, oxazolidinone or acetanilide. **b**, *In situ* observation of an organo-Cu(III) intermediate in the

coupling of aryl bromide **12** with pyridone. Tf, trifluoromethanesulphonate. **c**, *In situ* observation of an organo-Cu(III) intermediate in the oxygenation of carbon–hydrogen bonds. Me, methyl.

The formation of carbon–halogen bonds has been an important target reaction for catalysis by high-valent palladium. Halogenated molecules are valuable starting materials for many organic transformations including nucleophilic substitutions, metal-catalysed cross-couplings and Friedel–Crafts alkylations. Notably, carbon–halogen bond-forming reductive elimination is both thermodynamically unfavourable and kinetically slow from most Pd(II) complexes⁶⁰. As a result, most Pd(II)- or Pd(0)-catalysed transformations of aryl or alkyl halides involve breaking carbon–halogen bonds rather than forming them⁶¹ (Fig. 5a). By marked contrast, recent work has shown that many high-valent palladium complexes promote the facile formation of carbon–halogen bonds^{62–64}. Initial studies in this area focused on generating high-valent organometallic palladium halide complexes through the stoichiometric two-electron oxidation of Pd(II) precursors with electrophilic halogenating reagents^{65–67} (for example Cl₂, PhICl₂, *N*-chlorosuccinimide and XeF₂ (collectively abbreviated X⁺ in Fig. 5a)). Depending on the structure of the Pd(II) starting material, these reactions afford either monomeric Pd(IV) complexes such as **19** (ref. 68), **20** (ref. 69) and **21** (ref. 70) or dimeric Pd(III) species such as **22** (ref. 71) (Fig. 4). Many of these high-valent palladium compounds are isolatable at room temperature (25 °C). However, on being heated, they all undergo kinetically fast and highly thermodynamically favourable carbon–heteroatom bond-forming reductive elimination to release halogenated organic products (Fig. 5a).

The stoichiometric studies shown in Fig. 5a have informed the development of new palladium-catalysed halogenation reactions that involve

high-valent intermediates. One particularly well-studied example involves the ligand-directed halogenation of arene and alkane carbon–hydrogen bonds^{62–64,72,73} (Fig. 5b). Electrophilic halogenating reagents (X⁺) are used to promote the formation of high-valent palladium intermediates during

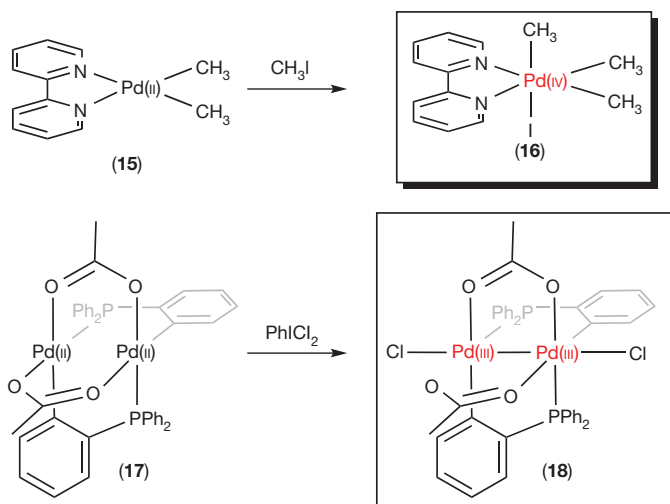


Figure 4 | Early examples of Pd(III) and Pd(IV) organometallic complexes. Ph, phenyl.

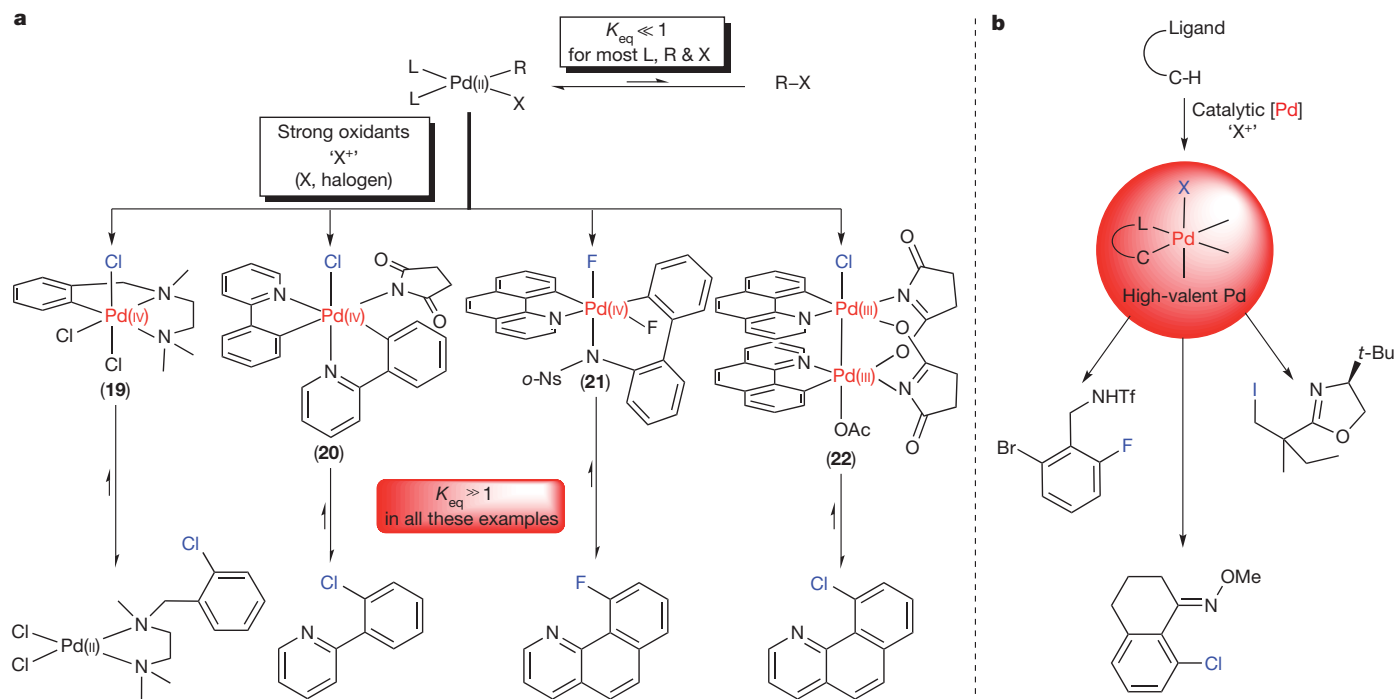


Figure 5 | High-valent palladium complexes involved in carbon-halogen bond formation. **a**, Carbon-halogen bond-forming reductive elimination is thermodynamically unfavourable from most Pd(II) species but not from high-valent palladium complexes such as **19–22**. Ac, acetyl; K_{eq} , equilibrium

constant. **b**, Select examples of palladium-catalysed carbon-hydrogen halogenation reactions. L, L-type ligand; R, aryl; X, halogen; *o*-Ns, *ortho*-nosyl; *t*-Bu, *tert*-butyl.

catalysis. Depending on the structure of the reagent, diverse carbon-halogen bonds can be formed. For example, *N*-fluoropyridinium salts generate carbon-fluorine bonds, *N*-chlorosuccinimide promotes the formation of carbon-chlorine bonds and acetyl hypoiodite provides access to iodinated products.

Detailed mechanistic studies of the palladium-catalysed carbon-hydrogen chlorination of benzo[*h*]quinoline with *N*-chlorosuccinimide support the intermediacy of a high-valent palladium species⁷¹ (Fig. 6a). The resting state of the catalyst was determined to be the dinuclear succinate-bridged Pd(II) complex **23** (Fig. 6b). This compound is a kinetically competent catalyst in the presence of added acetate. Furthermore, rate studies of the **23**-catalysed carbon-hydrogen chlorination reaction show a first-order dependence on this Pd(II)–Pd(II) dimer and a first-order dependence on the oxidant. These data are consistent with rate-limiting two-electron oxidation of **23** to generate the high-valent Pd(III)–Pd(III) dimer **22** (Fig. 6b). This Pd(III)–Pd(III) intermediate could not be observed under the catalytic conditions, as to be expected when oxidation is rate limiting. However, dimer **22** could be synthesized independently at -78°C . On being warmed to 23°C , **22** underwent carbon-chlorine bond-forming reductive elimination to release chlorinated product **24** in 84% yield, further supporting its intermediacy in catalysis.

High-valent palladium catalysis has also been exploited for the halofunctionalization of alkenes^{74–76} (that is, the addition of a halogen and another functional group), as exemplified by Fig. 6c. The combination of a palladium catalyst, an alkene and an aryl–metal reagent such as Bu_3SnPh is well known to produce a Pd(II)–alkyl intermediate like **25** (Fig. 6c). However, the fate of this intermediate and the ultimate organic product of the reaction vary drastically depending on the choice of oxidant. For example, with oxidants such as dioxygen (which has low kinetic reactivity with most Pd(II) complexes), **25** undergoes β -hydride elimination to release styrene product **26** through a conventional low-valent Pd(II) or Pd(0) manifold⁷⁷ (Fig. 6c, reaction (i)). By contrast, kinetically reactive Cl^+ oxidants such as PhICl_2 can rapidly intercept **25** to generate putative high-valent palladium intermediates. These can

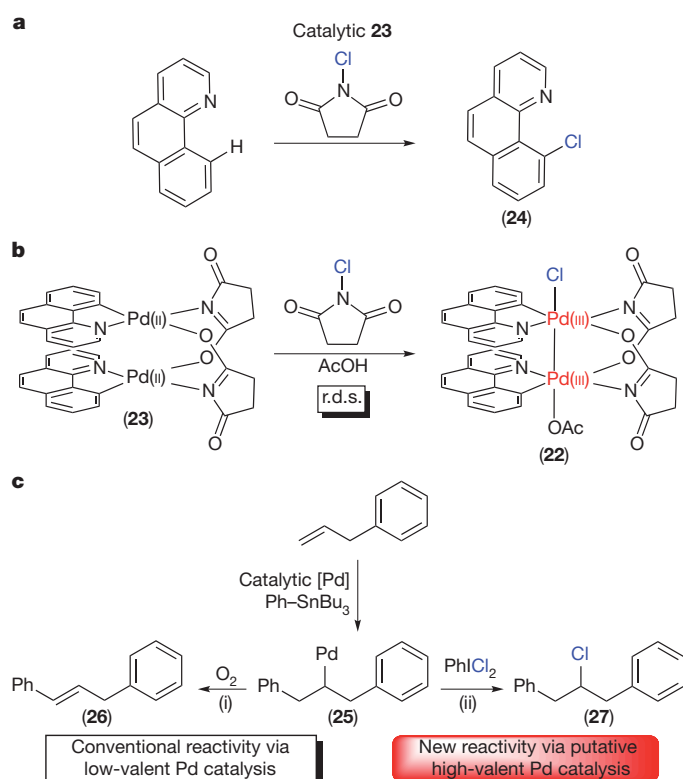


Figure 6 | Palladium-catalysed chlorination. **a**, Complex **23** is an efficient catalyst for the palladium-catalysed carbon-hydrogen chlorination of benzo[*h*]quinoline. **b**, The rate-determining step (r.d.s.) of this palladium-catalysed reaction is oxidation of **23** by *N*-chlorosuccinimide to form the Pd(III)–Pd(III) dimer **22**. **c**, High-valent palladium-catalysed 1,2-arylchlorination (ii) is complementary to low-valent palladium-catalysed reactions (i) of α -alkenes.

then undergo carbon–chlorine bond-forming reductive elimination to release 1,2-arylchlorinated compound **27** (ref. 76; Fig. 6c, reaction (ii)). Although the intermediacy of Pd(III) and/or Pd(IV) in these arylchlorination reactions has not yet been definitively confirmed, the observed reactivity (favouring carbon–chlorine bond formation over β -hydride elimination) is consistent with such a mechanism. This highlights another key complementarity between catalysis by low-valent palladium (where square-planar Pd(II)–alkyl intermediates typically undergo fast decomposition through β -hydride elimination) and catalysis by high-valent palladium^{62–64,72} (where β -hydride elimination is disfavoured owing to the lack of open coordination sites at octahedral Pd(III)– and/or Pd(IV)–alkyl complexes). Notably, many related palladium-catalysed alkene difunctionalization reactions have been reported over the past six years that also probably proceed through high-valent palladium pathways^{63,64}.

Another challenging and desirable chemical target for catalysis by high-valent palladium has been the generation of carbon–CF₃ linkages. Trifluoromethyl groups appear in numerous commercial pharmaceuticals and drug candidates and can drastically enhance the metabolic stability and bioavailability of biologically active molecules⁷⁸. Despite the prevalence of these groups in medicinal chemistry, efficient approaches to introducing CF₃ into organic compounds under mild conditions are limited. Methods involving metal catalysts are particularly rare^{79,80} and would be powerful synthetic tools to complement currently available chemical processes.

Many previous efforts to develop catalytic trifluoromethylation reactions have been hampered by the kinetic inertness of most metal–CF₃ complexes to carbon–CF₃ bond formation⁷⁹. For example, carbon–CF₃ coupling at Pd(II) centres requires specialized phosphine ligands to proceed efficiently^{79,80} (Fig. 7a). By contrast, Pd(IV) complexes containing simple bidentate, nitrogen donor ligands (N~N) undergo facile carbon–CF₃ bond-forming reductive elimination. For example, the stoichiometric reaction of (N~N)Pd(II)(aryl)CF₃ complexes with *N*-fluoropyridinium oxidants affords isolatable high-valent

(N~N)Pd(IV)(aryl)CF₃ intermediates. These compounds participate in rapid carbon–CF₃ coupling at temperatures as low as 25 °C (ref. 81).

A related approach has been used to achieve catalytic ligand-directed trifluoromethylation of aromatic carbon–hydrogen bonds. In this system, electrophilic trifluoromethylating reagents (CF₃⁺) were used to promote the formation of high-valent palladium intermediates, which decompose to afford aryl-trifluoromethylated products⁸². Remarkably, these reactions proceed efficiently with simple palladium salts as catalysts, and no external ligands (other than substrate) are required. Subsequent mechanistic studies suggested that the Pd(IV) complex **28** might be a catalytic intermediate⁸³ (Fig. 7b), as it serves as a kinetically competent catalyst under the reaction conditions. This methodology represents a transformation (ligand-directed C–H \rightarrow C–CF₃ conversion) that is not at present accessible using any other transition-metal catalyst, again highlighting the power of high-valent palladium to do novel chemistry.

In addition to the carbon–halogen and carbon–trifluoromethyl bond-forming reactions discussed above, high-valent palladium intermediates have also been implicated in the selective transformation of alkane and arene carbon–hydrogen bonds into carbon–oxygen, carbon–carbon, carbon–nitrogen and carbon–sulphur linkages. Detailed mechanistic investigations of catalytic carbon–hydrogen acetoxylation⁸⁴ and arylation³¹ have provided evidence consistent with the formation of high-valent palladium intermediates in these reactions as well.

There has been much recent progress in high-valent palladium catalysis. Over the past decade, numerous organometallic Pd(IV) and Pd(III) complexes have been synthesized by the reaction of Pd(II) starting materials with strong oxidants. A wide range of carbon–carbon and carbon–heteroatom bond-forming reductive elimination reactions can be achieved from these species, and the selectivity, reactivity and mechanisms of these transformations have been studied in detail. Furthermore, a number of these species have been shown to be kinetically competent catalysts for carbon–hydrogen bond halogenation, trifluoromethylation and other reactions. These results have firmly established the feasibility

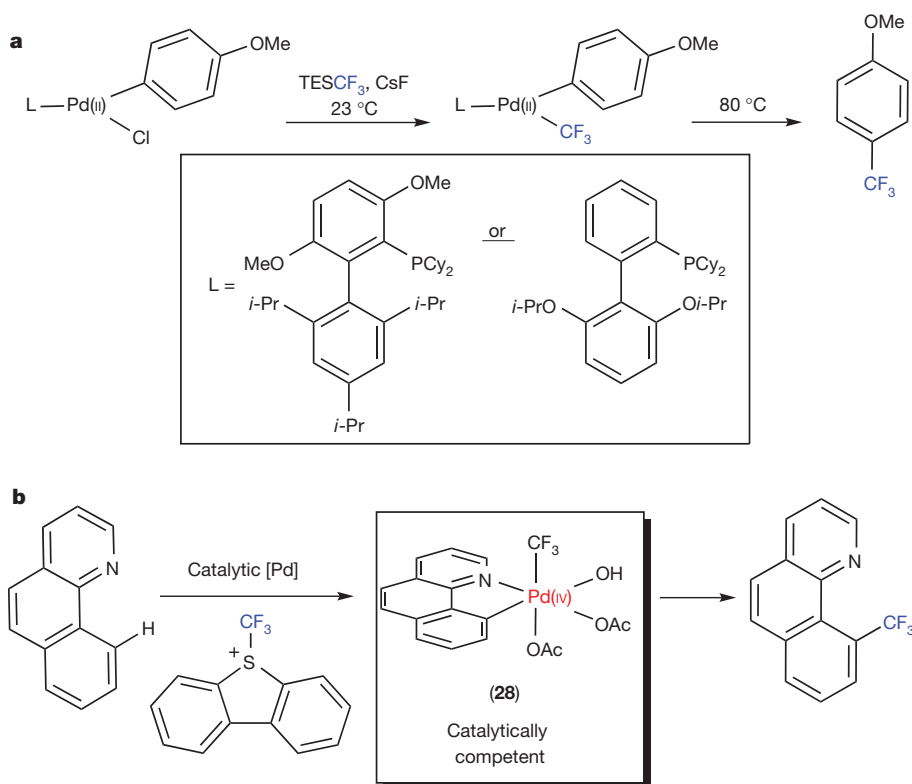


Figure 7 | Palladium-catalysed trifluoromethylation. **a**, Carbon–CF₃ bond formation from Pd(II) requires specialized phosphine ligands. TES, triethylsilyl; Cy, cyclohexyl; *i*-Pr = *iso*-propyl. **b**, Using a high-valent palladium strategy,

catalytic carbon–hydrogen trifluoromethylation has been developed through putative Pd(IV) intermediate **28**.

and synthetic utility of high-oxidation-state palladium catalysis in organic synthesis, and the field shows great potential for the future.

Comparison and contrast

As discussed here, high-valent copper chemistry has several features in common with high-valent palladium chemistry. First, although for each metal the specific transformations that have been selected for detailed mechanistic investigation vary significantly, Cu(III) and Pd(III)/Pd(IV) species have been shown to participate in closely related carbon–carbon and carbon–heteroatom bond-forming reductive elimination reactions. One particularly striking example is the accessibility of carbon–halogen bond-forming reductive elimination from Cu(III) complex **13** as well as from Pd(IV)/Pd(III) complexes **19–22**. Similar ligand environments have been shown to stabilize high-valent complexes of both metals. In particular, rigid multidentate ligands (such as the macrocycles of copper complexes **2**, **10** and **13** and the cyclometalated benzo[*h*]quinoline of Pd(IV) species **21** and **28** and Pd(III) complex **22**) tend to slow competing reductive elimination processes. Furthermore, the presence of multiple highly electron-donating σ -aryl or σ -alkyl ligands (as in Cu(III) complexes **4–9** and Pd(IV) complexes **16** and **19–21**) facilitate the detection or isolation of high-valent species of both copper and palladium.

Recent examples of copper- or palladium-catalysed oxidation reactions reveal additional intriguing similarities. There are a multitude of catalytic carbon–carbon and carbon–heteroatom coupling reactions that share the following features: a copper or palladium catalyst, an oxidant and an organic substrate that is a precursor to a metal–carbon bond (such as an aryl halide, a carbon–hydrogen bond or a transmetalating reagent). Three examples of such transformations are shown in Fig. 8. Although detailed mechanistic analysis will be required to establish firmly the pathway for each system, it seems likely that many (if not all) of these reactions proceed through high-valent copper or palladium manifolds. In the first example, copper and palladium catalyse the same overall reaction, the ligand-directed carbon–hydrogen acetoxylation of 2-phenylpyridine^{52,85} (Fig. 8a). In the second example, the same oxidant (S-(trifluoromethyl) dibenzothiophenium) is used to effect the trifluoromethylation of two different organic substrates^{82,86} (Fig. 8b). Finally, in the third example, both metals catalyse the carbon–hydrogen arylation of indole with diaryliodonium salts^{87,88} (Fig. 8c).

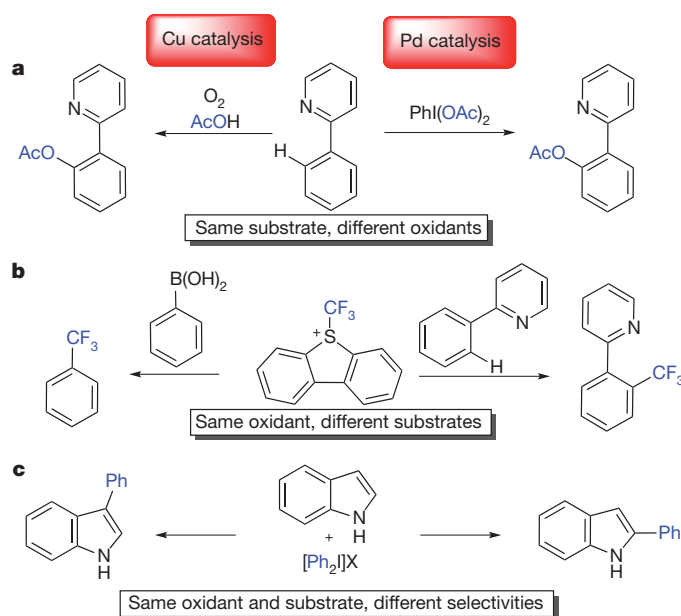


Figure 8 | Oxidative bond-forming reactions catalysed by copper and palladium. These reactions exemplify similarities and differences between copper and palladium.

The reactions in Fig. 8 not only illustrate key similarities but also highlight key differences or complementarities between the oxidative chemistry of palladium and that of copper. For example, the palladium-catalysed carbon–hydrogen acetoxylation of 2-phenylpyridine (Fig. 8a) requires the use of $\text{PhI}(\text{OAc})_2$ as the terminal oxidant⁸⁵. This reagent is quite expensive and generates an equivalent of iodobenzene waste with each catalytic turnover. By marked contrast, the copper-catalysed acetoxylation uses abundant and atom-economical dioxygen as the oxidant⁵². The ability to generate high-valent copper using dioxygen is currently a distinct advantage of high-valent copper catalysis^{40,42,53}. Although dioxygen is thermodynamically capable of oxidizing Pd(II) to Pd(IV), most organo-Pd(II) intermediates are kinetically inert to oxidation by dioxygen⁷³. However, two recent reports have shown that aerobic, palladium-catalysed, ligand-directed carbon–hydrogen oxygenation is possible (potentially through high-valent palladium intermediates), suggesting a promising future in this area^{89,90}.

As shown in Fig. 8b, both copper and palladium catalyse carbon– CF_3 bond-forming reactions with S-(trifluoromethyl) dibenzothiophenium^{82,86}. Catalytically competent Pd(IV) intermediates have been observed and isolated in the carbon–hydrogen trifluoromethylation reaction⁸³ (Fig. 7b). By contrast, mechanisms involving Cu(III)(aryl) CF_3 intermediates have been proposed but remain to be confirmed experimentally for the copper-catalysed trifluoromethylation of boronic acids⁸⁶. These two examples demonstrate another key complementarity between palladium and copper catalysis. High-valent palladium catalysis has been used to transform carbon–hydrogen substrates into many different functional groups (with trifluoromethyl being just one example), and palladium-catalysed carbon–hydrogen oxidation is an extremely common, general, well-studied reaction⁷³. By marked contrast, high-valent copper catalysis has predominantly focused on prefunctionalized substrates such as aryl boronic acids and aryl halides (Fig. 8b). At present, copper-catalysed carbon–hydrogen bond oxidation reactions are comparatively rare and the range of possible substrates is significantly narrower than that of analogous palladium-catalysed reactions^{50–53}. For example, whereas the copper-catalysed functionalization of unactivated alkane carbon–hydrogen bonds remains highly challenging, such transformations are increasingly common using palladium⁷³. The development of more-robust and general methods for carbon–hydrogen bond oxidation through catalysis by high-valent copper is likely to be a major thrust of research in this field.

Finally, both copper and palladium catalyse the arylation of indole with diaryliodonium salts; however, the site selectivities of these two reactions are orthogonal (Fig. 8c). Whereas the palladium-catalysed reaction results in selective arylation at the 2-position⁸⁷, the copper-catalysed methods can be tuned to give exclusive arylation at the 3-position⁸⁸. Site selectivity is one of the most difficult challenges in the field of carbon–hydrogen functionalization. As such, the ability to tune selectivity as a function of the metal is of great potential synthetic utility. Thus far, neither of these transformations has been the subject of detailed mechanistic analysis, but the generation of Cu(III) and Pd(IV) intermediates has been suggested in both cases. In combination, the examples in Fig. 8 demonstrate the tremendous opportunities available in the concurrent development of high-valent copper and palladium catalysis.

Looking forward

The fields of high-valent palladium and copper chemistry are sure to have a bright and rapidly expanding future. It will be critical to increase our understanding of, and enhance the chemo-, regio- and stereoselectivity of, catalytic processes involving high-valent copper and palladium intermediates. In many cases, the coordination sphere of these high-valent metal centres contains multiple possible partners for reductive bond-forming reactions. The ability to control the chemoselectivity of the bond-forming event is of central importance in achieving efficient, high-yield catalytic transformations. In addition, the identification of chiral ligands that are compatible with catalysis by high-valent copper

and/or palladium could potentially allow novel asymmetric conjugate addition, aryl–aryl coupling, carbon–hydrogen oxidation and/or alkene difunctionalization reactions, which would all be of great value for organic synthesis.

Additional future work in this field will focus on expanding the scope of the fundamental organometallic reactions that are possible at high-valent copper and palladium centres. Despite the impressive progress described above, the synthetic power of high-valent organometallic intermediates has thus far been explored quite narrowly, with an almost exclusive focus on reductive bond-forming reactions. We anticipate that the design of new ancillary ligands that even better stabilize high-valent palladium and copper will facilitate the study and application of carbon–hydrogen activation, σ -bond metathesis, migratory insertion and nucleopalladation reactions at these metal centres. Such reactions could potentially proceed with novel patterns of reactivity and selectivity relative to analogous transformations of low-valent analogues. For example, several preliminary reports have suggested that carbon–hydrogen activation occurs with completely different site selectivities at Pd(IV) centres than at Pd(II) centres^{91–94}.

Finally, a number of recent reports suggest that high-valent organometallic complexes of other late transition metals can catalyse reactions similar to those discussed for Cu(III) and Pd(III)/Pd(IV) above. For example, complexes of nickel(III), nickel(IV), silver(II) and silver(III) have been observed and/or implicated in carbon–halogen and carbon–nitrogen bond-forming processes^{95–100}. Further exploration of these is likely to uncover many additional applications for high-valent late transition metals in catalysis.

- Magano, J. & Dunetz, J. R. Large-scale applications of transition metal-catalyzed couplings for the synthesis of pharmaceuticals. *Chem. Rev.* **111**, 2177–2250 (2011).
- Evano, G., Blanchard, N. & Toumi, M. Copper-mediated coupling reactions and their applications in natural products and designed biomolecules synthesis. *Chem. Rev.* **108**, 3054–3131 (2008).
- Krause, N. (ed.) *Modern Organocopper Chemistry* (Wiley-VCH, 2002).
- Eckert, M., Fleischmann, G., Jira, R., Bolt, H. M. & Golka, K. in *Ullmann's Encyclopedia of Industrial Chemistry* 191–207 (Wiley-VCH, 2006).
- Corbet, J.-P. & Mignani, G. Selected patented cross-coupling reaction technologies. *Chem. Rev.* **106**, 2651–2710 (2006).
- Ullmann, F. On a new formation of diphenylamine derivatives. *Ber. Deutsch. Chem. Ges.* **36**, 2382–2384 (1903).
- Ullmann, F. Over a new preparation manner of phenylethersalicylic acid. *Ber. Deutsch. Chem. Ges.* **37**, 853–854 (1904).
- Beletskaya, I. P. & Chepravkov, A. V. Copper in cross-coupling reactions: the post-Ullmann chemistry. *Coord. Chem. Rev.* **248**, 2337–2364 (2004).
- Monnier, F. & Taillefer, M. Catalytic C–C, C–N, and C–O Ullmann-type coupling reactions. *Angew. Chem. Int. Ed.* **48**, 6954–6971 (2009).
- Ribas, X. & Casitas, A. in *Ideas in Chemistry and Molecular Sciences: Where Chemistry Meets Life* (ed. Pignataro, B.) 31–57 (Wiley-VCH, 2010).
- This is a thorough review on high-valent copper chemistry.
- Sperotto, E., van Klink, G. P. M., van Koten, G. & de Vries, J. G. The mechanism of the modified Ullmann reaction. *Dalton Trans.* **39**, 10338–10351 (2010).
- Nakamura, E. & Mori, S. Wherefore art thou copper? Structures and reaction mechanisms of organocuprate clusters in organic chemistry. *Angew. Chem. Int. Ed.* **39**, 3750–3771 (2000).
- Torborg, C. & Beller, M. Recent applications of palladium-catalyzed coupling reactions in the pharmaceutical, agrochemical, and fine chemical industries. *Adv. Synth. Catal.* **351**, 3027–3043 (2009).
- Hartwig, J. F. Electronic effects on reductive elimination to form carbon–carbon and carbon–heteroatom bonds from palladium(II) complexes. *Inorg. Chem.* **46**, 1936–1947 (2007).
- Willert-Porada, M. A., Burton, D. J. & Baenziger, N. C. Synthesis and X-ray structure of bis(trifluoromethyl)(N,N-diethyldithiocarbamate)-copper; a remarkably stable perfluoroalkylcopper(III) complex. *J. Chem. Soc. Chem. Commun.* 1633–1634 (1989).
- Furuta, H., Maeda, H. & Osuka, A. Doubly *N*-confused porphyrin: a new complexing agent capable of stabilizing higher oxidation states. *J. Am. Chem. Soc.* **122**, 803–807 (2000).
- Santo, R. et al. Diamagnetic-paramagnetic conversion of tris(2-pyridylthio)methylcopper(III) through a structural change from trigonal bipyramidal to octahedral. *Angew. Chem. Int. Ed.* **45**, 7611–7614 (2006).
- Cohen, T., Wood, J. & Dietz, A. G. Organocopper intermediates in the exchange reaction of aryl halides with salts of copper(I). The possible role of copper(III). *Tetrahedr. Lett.* **15**, 3555–3558 (1974).
- Dorigo, A. E., Wanner, J. & Schleyer, P. R. Computational evidence for the existence of Cu^{III} intermediates in addition and substitution reactions with dialkylcuprates. *Angew. Chem. Int. Ed. Engl.* **34**, 476–478 (1995).
- Snyder, J. P. Mechanism of lithium cuprate conjugate addition: neutral tetracoordinate Cu^I cuprates as essential intermediates. *J. Am. Chem. Soc.* **117**, 11025–11026 (1995).
- Karlström, A. S. E. & Bäckvall, J.-E. Experimental evidence supporting a Cu^{III} intermediate in cross-coupling reactions of allylic esters with dialkylcuprate species. *Chemistry* **7**, 1981–1989 (2001).
- Bertz, S. H., Cope, S., Murphy, M., Ogle, C. A. & Taylor, B. J. Rapid injection NMR in mechanistic organocopper chemistry. Preparation of the elusive copper(III) intermediate. *J. Am. Chem. Soc.* **129**, 7208–7209 (2007).
- This paper reports the first observation of Cu(III) intermediates in conjugate addition reactions using RI-NMR.
- Hu, H. & Snyder, J. P. Organocuprate conjugate addition: the square-planar “Cu^{III}” intermediate. *J. Am. Chem. Soc.* **129**, 7210–7211 (2007).
- Bertz, S. H., Cope, S., Dorton, D., Murphy, M. & Ogle, C. A. Organocuprate cross-coupling: the central role of the copper(III) intermediate and the importance of the copper(I) precursor. *Angew. Chem. Int. Ed.* **46**, 7082–7085 (2007).
- Gärtner, T., Henze, W. & Gschwind, R. M. NMR detection of Cu(III) intermediates in substitution reactions of alkyl halides with Gilman cuprates. *J. Am. Chem. Soc.* **129**, 11362–11363 (2007).
- Bartholomew, E. R., Bertz, S. H., Cope, S., Murphy, M. & Ogle, C. A. Preparation of σ - and π -allylcopper(III) intermediates in S_N2 and S_N2' reactions of organocuprate(I) reagents with allylic substrates. *J. Am. Chem. Soc.* **130**, 11244–11245 (2008).
- Bertz, S. H., Miao, G. & Eriksson, M. It's on lithium! An answer to the recent communication which asked the question: “If the cyano ligand is not on copper, then where is it?” *Chem. Commun. (Camb.)* 815–816 (1996).
- Alexakis, A., Vastra, J. & Mangeney, P. Acceleration of the conjugate addition of diethyl zinc to enones by either Cu(OTf)₂ or trivalent phosphorous ligands. *Tetrahedr. Lett.* **38**, 7745–7748 (1997).
- Bartholomew, E. R. et al. Neutral organocopper(III) complexes. *Chem. Commun. (Camb.)* 1176–1177 (2008).
- Phipps, R. J. & Gaunt, M. J. A meta-selective copper-catalyzed C–H bond arylation. *Science* **323**, 1593–1597 (2009).
- Deprez, N. R. & Sanford, M. S. Synthetic and mechanistic studies of Pd-catalyzed C–H arylation with diaryliodonium salts: evidence for a bimetallic high oxidation state Pd intermediate. *J. Am. Chem. Soc.* **131**, 11234–11241 (2009).
- Chen, B., Hou, X.-L., Li, Y.-X. & Wu, Y.-D. Mechanistic understanding of the unexpected meta selectivity in copper-catalyzed anilide C–H bond arylation. *J. Am. Chem. Soc.* **133**, 7668–7671 (2011).
- Das, P., Sharma, D., Kumar, M. & Singh, B. Copper promoted C–N and C–O type cross-coupling reactions. *Curr. Org. Chem.* **14**, 754–783 (2010).
- Chan, D. M. T. & Lam, P. Y. S. in *Boronic Acids* (ed. Hall, D. G.) 205–240 (Wiley-VCH, 2005).
- Casitas, A. et al. Direct observation of Cu^I/Cu^{III} redox steps relevant to Ullmann-type coupling reactions. *Chem. Sci.* **1**, 326–330 (2010).
- Kunz, K., Scholz, U. & Ganzer, D. Renaissance of Ullmann and Goldberg reactions: progress in copper catalyzed C–N, C–O, and C–S-coupling. *Synlett* **15**, 2428–2439 (2003).
- King, A. E. et al. Copper-catalyzed aerobic oxidative functionalization of an arene C–H bond: evidence for an aryl-copper(III) intermediate. *J. Am. Chem. Soc.* **132**, 12068–12073 (2010).
- This paper demonstrates a Cu(III)–aryl complex to be a catalytically relevant intermediate in a carbon–hydrogen bond amination and methoxylation reaction.
- Brasche, G. & Buchwald, S. L. C–H functionalization/C–N bond formation: copper-catalyzed synthesis of benzimidazoles from amidines. *Angew. Chem. Int. Ed.* **47**, 1932–1934 (2008).
- Ueda, S. & Nagasawa, H. Synthesis of 2-arylbenzoxazoles by copper-catalyzed intramolecular oxidative C–O coupling of benzanilides. *Angew. Chem. Int. Ed.* **47**, 6411–6413 (2008).
- Hamada, T., Ye, X. & Stahl, S. S. Copper-catalyzed aerobic oxidative amidation of terminal alkynes: efficient synthesis of ynamides. *J. Am. Chem. Soc.* **130**, 833–835 (2008).
- Uemura, T., Imoto, S. & Chatani, N. Amination of the ortho C–H bonds by the Cu(OAc)₂-mediated reaction of 2-phenylpyridines with anilines. *Chem. Lett.* **35**, 842–843 (2006).
- Mizuhara, T., Inuki, S., Oishi, S., Fujii, N. & Ohno, H. Cu(II)-mediated oxidative intermolecular ortho C–H functionalisation using tetrahydropyrimidine as the directing group. *Chem. Commun. (Camb.)* 3413–3415 (2009).
- Huffman, L. M. & Stahl, S. S. Carbon–nitrogen bond formation involving well-defined aryl-copper(II) complexes. *J. Am. Chem. Soc.* **130**, 9196–9197 (2008).
- Tye, J. W., Weng, Z., Johns, A. M., Incarvito, C. D. & Hartwig, J. F. Copper complexes of anionic nitrogen ligands in the amidation and imidation of aryl halides. *J. Am. Chem. Soc.* **130**, 9971–9983 (2008).
- Jones, G. O., Liu, P., Houk, K. N. & Buchwald, S. L. Computational explorations of mechanisms and ligand-directed selectivities of copper-catalyzed Ullmann-type reactions. *J. Am. Chem. Soc.* **132**, 6205–6213 (2010).
- Yu, H.-Z., Jiang, Y.-Y., Fu, Y. & Liu, L. Alternative mechanistic explanation for ligand-dependent selectivities in copper-catalyzed *N*- and *O*-arylation reactions. *J. Am. Chem. Soc.* **132**, 18078–18091 (2010).
- Ribas, X. et al. Aryl C–H activation by Cu^{II} to form an organometallic aryl-Cu^{III} species: a novel twist on copper disproportionation. *Angew. Chem. Int. Ed.* **41**, 2991–2994 (2002).
- This paper reports an early example of an isolatable, well-characterized Cu(III)–aryl complex.
- Wendlandt, A. E., Suess, A. M. & Stahl, S. S. Copper-catalyzed aerobic oxidative C–H functionalizations: trends and mechanistic insights. *Angew. Chem. Int. Ed.* **50**, 11062–11087 (2011).

49. Ueda, S. & Nagasawa, H. Copper-catalyzed synthesis of benzoxazoles via a regioselective C-H functionalization/C-O bond formation under an air atmosphere. *J. Org. Chem.* **74**, 4272–4277 (2009).
50. Siemsen, P., Livingston, R. C. & Diederich, F. Acetylenic coupling: a powerful tool in molecular construction. *Angew. Chem. Int. Ed.* **39**, 2632–2657 (2000).
51. Gao, Y. *et al.* Copper-catalyzed aerobic oxidative coupling of terminal alkynes with *H*-phosphates leading to alkynylphosphonates. *J. Am. Chem. Soc.* **131**, 7956–7957 (2009).
52. Chen, X., Hao, X.-S., Goodhue, C. E. & Yu, J.-Q. Cu(II)-catalyzed functionalizations of aryl C–H bonds using O₂ as an oxidant. *J. Am. Chem. Soc.* **128**, 6790–6791 (2006).
53. Wang, W., Luo, F., Zhang, S. & Cheng, J. Copper(II)-catalyzed ortho-acyloxylation of the 2-arylpyridines sp² C–H bonds with anhydrides, using O₂ as terminal oxidant. *J. Org. Chem.* **75**, 2415–2418 (2010).
54. Casitas, A., Canta, M., Solá, M., Costas, M. & Ribas, X. Nucleophilic aryl fluorination and aryl halide exchange mediated by a Cu^I/Cu^{III} catalytic cycle. *J. Am. Chem. Soc.* **133**, 19386–19392 (2011).
55. Heck, R. F. Aromatic haloethylation with palladium and copper halides. *J. Am. Chem. Soc.* **90**, 5538–5542 (1968).
56. Fahey, D. R. The coordination-catalyzed ortho-halogenation of azobenzene. *J. Organomet. Chem.* **27**, 283–292 (1971).
57. Tremont, S. J. & Rahman, H. U. Ortho-alkylation of acetanilides using alkyl halides and palladium acetate. *J. Am. Chem. Soc.* **106**, 5759–5760 (1984).
58. Byers, P. K., Canty, A. J., Skelton, B. W. & White, A. H. Oxidative addition of iodomethane to [PdMe₂(bpy)] and the X-ray structure of the organopalladium(IV) product fac-[PdMe₃(bpy)] (bpy = 2,2'-bipyridyl). *J. Chem. Soc. Chem. Commun.* 1722–1724 (1986).
59. Cotton, F. A. *et al.* High yield syntheses of stable, singly bonded Pd₂⁶⁺ compounds. *J. Am. Chem. Soc.* **128**, 13674–13675 (2006).
60. Roy, A. H. & Hartwig, J. F. Directly observed reductive elimination of aryl halides from monomeric arylpalladium(II) halide complexes. *J. Am. Chem. Soc.* **125**, 13944–13945 (2003).
61. Collman, J. P., Hegedus, L. S., Norton, J. R. & Finke, R. G. *Principles and Applications of Organotransition Metal Chemistry* 2nd edn 322–333 (University Science Books, 1982).
62. Canty, A. J. Organopalladium and platinum chemistry in oxidizing milieu as models for organic synthesis involving the higher oxidation states of palladium. *Dalton Trans.* **47**, 10409–10417 (2009).
63. Muñoz, K. High-oxidation-state palladium catalysis: new reactivity for organic synthesis. *Angew. Chem. Int. Ed.* **48**, 9412–9423 (2009).
This is a thorough review on high-valent palladium chemistry.
64. Xu, L.-M., Li, B.-J., Yang, Z. & Shi, Z.-J. Organopalladium(IV) chemistry. *Chem. Soc. Rev.* **39**, 712–733 (2010).
65. Vicente, J., Arcas, A., Julia-Hernandez, F. & Bautista, D. Synthesis, isolation, and characterization of an organometallic triiodopalladium(IV) complex. Quantitative and regioselective synthesis of two C–I reductive elimination products. *Inorg. Chem.* **50**, 5339–5341 (2011).
66. Arnold, P. L., Sanford, M. S. & Pearson, S. M. Chelating *N*-heterocyclic carbene alkoxide as a supporting ligand for Pd^{IV} C–H bond functionalization catalysis. *J. Am. Chem. Soc.* **131**, 13912–13913 (2009).
67. Ball, N. D. & Sanford, M. S. Synthesis and reactivity of a mono-σ-aryl palladium(IV) fluoride complex. *J. Am. Chem. Soc.* **131**, 3796–3797 (2009).
68. Alsters, P. L. *et al.* Rigid five- and six-membered C,N,N'-bound aryl, benzyl, and alkyl organopalladium complexes: sp² vs sp³ C–H activation during cyclopalladation and palladium(IV) intermediates in oxidative addition reactions with dihalogens and alkyl halides. *Organometallics* **12**, 1831–1844 (1993).
69. Whitfield, S. R. & Sanford, M. S. Reactivity of Pd(II) complexes with electrophilic chlorinating reagents: isolation of Pd(IV) products and observation of C–Cl bond-forming reductive elimination. *J. Am. Chem. Soc.* **129**, 15142–15143 (2007).
70. Furuya, T. *et al.* Mechanism of C–F reductive elimination from palladium(IV) fluoride. *J. Am. Chem. Soc.* **132**, 3793–3807 (2010).
71. Powers, D. C., Xiao, D. Y., Geibel, M. A. L. & Ritter, T. On the mechanism of palladium-catalyzed aromatic C–H oxidation. *J. Am. Chem. Soc.* **132**, 14530–14536 (2010).
This paper describes mechanistic studies implicating a Pd(III) dimer intermediate in a palladium-catalysed carbon-hydrogen chlorination reaction described in ref. 85.
72. Sehna, P., Taylor, R. J. K. & Fairlamb, I. J. S. Emergence of palladium(IV) chemistry in synthesis and catalysis. *Chem. Rev.* **110**, 824–889 (2010).
73. Lyons, T. W. & Sanford, M. S. Palladium-catalyzed ligand-directed C–H functionalization reactions. *Chem. Rev.* **110**, 1147–1169 (2010).
This is a comprehensive review of palladium-catalysed, ligand-directed carbon-hydrogen functionalization reactions.
74. Wu, T., Yin, G. & Liu, G. Palladium-catalyzed intramolecular aminofluorination of unactivated alkenes. *J. Am. Chem. Soc.* **131**, 16354–16355 (2009).
75. Michael, F. E., Sibbald, P. A. & Cochran, B. M. Palladium-catalyzed intramolecular chloroamination of alkenes. *Org. Lett.* **10**, 793–796 (2008).
76. Kalyani, D., Satterfield, A. D. & Sanford, M. S. Palladium-catalyzed oxidative arylhalogenation of alkenes: synthetic scope and mechanistic insights. *J. Am. Chem. Soc.* **132**, 8419–8427 (2010).
This paper describes high-valent palladium-catalysed arylhalogenation reactions of α-alkenes and demonstrates complementarity to low-valent methods.
77. Oestreich, M. (ed.) *The Mizoroki-Heck Reaction* (Wiley, 2009).
78. Müller, K., Faeh, C. & Diederich, F. Fluorine in pharmaceuticals: looking beyond intuition. *Science* **317**, 1881–1886 (2007).
79. Grushin, V. V. The organometallic fluorine chemistry of palladium and rhodium: studies toward aromatic fluorination. *Acc. Chem. Res.* **43**, 160–171 (2010).
80. Cho, E. J. *et al.* The palladium-catalyzed trifluoromethylation of aryl chlorides. *Science* **328**, 1679–1681 (2010).
81. Ball, N. D., Gary, J. B., Ye, Y. & Sanford, M. S. Mechanistic and computational studies of oxidatively-induced bond-formation at Pd: rational design of room temperature trifluoromethylation. *J. Am. Chem. Soc.* **133**, 7577–7584 (2011).
82. Wang, X., Truesdale, L. & Yu, J.-Q. Pd(II)-catalyzed ortho-trifluoromethylation of arenes using TFA as a promoter. *J. Am. Chem. Soc.* **132**, 3648–3649 (2010).
This paper reports a novel carbon-hydrogen trifluoromethylation reaction catalysed by putative high-valent palladium.
83. Ye, Y., Ball, N. D., Kampf, J. W. & Sanford, M. S. Oxidation of catalytically relevant palladium dimer with “CF₃⁺”: formation and reactivity of a monomeric palladium(IV) aquo product. *J. Am. Chem. Soc.* **132**, 14682–14687 (2010).
This paper provides evidence supporting the possibility of a high-valent Pd(IV) intermediate in the carbon-hydrogen trifluoromethylation reaction described in ref. 82.
84. Powers, D. C., Geibel, M. A. L., Klein, J. E. M. N. & Ritter, T. Bimetallic palladium catalysis: direct observation of Pd(III)–Pd(II) intermediates. *J. Am. Chem. Soc.* **131**, 17050–17051 (2009).
85. Dick, A. R., Hull, K. L. & Sanford, M. S. A highly selective catalytic method for the oxidative functionalization of C–H bonds. *J. Am. Chem. Soc.* **126**, 2300–2301 (2004).
86. Xu, J. *et al.* Copper-catalyzed trifluoromethylation of aryl boronic acids using a CF₃⁺ reagent. *Chem. Commun. (Camb.)* **47**, 4300–4302 (2011).
87. Deprez, N. R., Kalyani, D., Krause, A. & Sanford, M. S. Room temperature palladium-catalyzed 2-arylation of indoles. *J. Am. Chem. Soc.* **128**, 4972–4973 (2006).
88. Phipps, R. J., Grimster, N. P. & Gaunt, M. J. Cu(II)-catalyzed direct and site-selective arylation of indoles under mild conditions. *J. Am. Chem. Soc.* **130**, 8172–8174 (2008).
89. Zhang, J., Khaskin, E., Anderson, N. P., Zavalij, P. Y. & Vedernikov, A. N. Catalytic aerobic oxidation of substituted 8-methylquinolines in Pd^{II}–2,6-pyridine dicarboxylic acid systems. *Chem. Commun. (Camb.)* 3625–3627 (2008).
90. Zhang, Y.-H. & Yu, J.-Q. Pd(II)-catalyzed hydroxylation of arenes with 1 atm O₂ or air. *J. Am. Chem. Soc.* **131**, 14654–14655 (2009).
91. Hull, K. L., Lanni, E. L. & Sanford, M. S. Highly regioselective catalytic oxidative coupling reactions: synthetic scope and mechanistic insights. *J. Am. Chem. Soc.* **128**, 14047–14049 (2006).
92. Rosewall, C. F., Sibbald, P. A., Liskin, D. V. & Michael, F. E. Palladium-catalyzed carboamination of alkenes promoted by *N*-fluorobenzenesulfonimide via C–H activation of arenes. *J. Am. Chem. Soc.* **131**, 9488–9489 (2009).
93. Hickman, A. J., & Sanford, M. S. Catalyst control of site selectivity in the Pd^{IV}/IV-catalyzed direct arylation of naphthalene. *ACS Catal.* **1**, 170–174 (2011).
94. Racowski, J. M., Ball, N. D. & Sanford, M. S. C–H bond activation at palladium(IV) centers. *J. Am. Chem. Soc.* **133**, 18022–18025 (2011).
95. Grove, D. M., van Koten, F., Zoet, R., Murrall, N. W. & Welch, A. J. Unique stable organometallic nickel(III) complexes; syntheses and the molecular structure of [Ni(C₆H₅(CH₂NMe₂)₂-2,6)]₂. *J. Am. Chem. Soc.* **105**, 1379–1380 (1983).
96. Ceder, R. M., Granell, J. & Muller, G. Preparation of five-membered nickelacycles with anionic C–N–N' terdentate ligands. X-ray crystal structure of [NiCl₂-(CH=NCH₂CH₂NMe₂)-3-ClC₆H₃]. *Organometallics* **15**, 4618–4624 (1996).
97. Higgs, A. T., Zinn, P. J., Simmons, S. J. & Sanford, M. S. Oxidatively induced carbon-halogen bond-forming reactions at nickel. *Organometallics* **28**, 6142–6144 (2009).
98. Muñoz, K. & Streuff, J. Exploring the nickel-catalyzed oxidation of alkenes: a diamination by sulfanamide transfer. *Angew. Chem. Int. Ed.* **46**, 7125–7127 (2007).
99. Lin, B. L., Clough, C. R. & Hillhouse, G. L. Interactions of aziridines with nickel complexes: oxidative-addition and reductive-elimination reactions that break and make C–N bonds. *J. Am. Chem. Soc.* **124**, 2890–2891 (2002).
100. Tang, P., Furuya, T., & Ritter, T. Silver-catalyzed late-stage fluorination. *J. Am. Chem. Soc.* **132**, 12150–12154 (2010).

Acknowledgements The work of the group of M.S.S. described herein was supported by the NIH-NIGMS (GM073836) and by the US NSF (CHE-0545909 and CHE-1111563).

Author Contributions. A.J.H. and M.S.S. worked together to outline the content, as well as write and edit the manuscript, references and figures.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to M.S.S. (mssanfor@umich.edu).

Emerging fungal threats to animal, plant and ecosystem health

Matthew C. Fisher¹, Daniel A. Henk¹, Cheryl J. Briggs², John S. Brownstein³, Lawrence C. Madoff⁴, Sarah L. McCraw⁵ & Sarah J. Gurr⁵

The past two decades have seen an increasing number of virulent infectious diseases in natural populations and managed landscapes. In both animals and plants, an unprecedented number of fungal and fungal-like diseases have recently caused some of the most severe die-offs and extinctions ever witnessed in wild species, and are jeopardizing food security. Human activity is intensifying fungal disease dispersal by modifying natural environments and thus creating new opportunities for evolution. We argue that nascent fungal infections will cause increasing attrition of biodiversity, with wider implications for human and ecosystem health, unless steps are taken to tighten biosecurity worldwide.

Emerging infectious diseases (EIDs) caused by fungi are increasingly recognized as presenting a worldwide threat to food security^{1,2} (Table 1 and Supplementary Table 1). This is not a new problem and fungi have long been known to constitute a widespread threat to plant species. Plant disease epidemics caused by fungi and the fungal-like oomycetes have altered the course of human history. In the nineteenth century, late blight led to starvation, economic ruin and the downfall of the English government during the Irish potato famine and, in the twentieth century, Dutch elm blight and chestnut blight laid bare urban and forest landscapes. The threat of plant disease has not abated, in fact it is heightened by resource-rich farming practices and exaggerated in the landscape by microbial adaptation to new ecosystems, brought about by trade and transportation³, and by climate fluctuations^{4,5}.

However, pathogenic fungi (also known as mycoses) have not been widely recognized as posing major threats to animal health. This perception is changing rapidly owing to the recent occurrence of several high-profile declines in wildlife caused by the emergence of previously unknown fungi^{6,7}. For example, during March 2007, a routine census of bats hibernating in New York State revealed mass mortalities⁸. Within a group of closely clustered caves, four species of bats were marked by a striking fungus growing on their muzzles and wing membranes, and the name 'white nose syndrome' (WNS) was coined. After the initial outbreak, the ascomycete fungus *Geomyces destructans* was shown to fulfil Koch's postulates and was described as the cause of WNS in American bat species^{9,10}. Mortalities exhibiting WNS have subsequently been found in an increasing number of bat overwintering sites and, by 2010, the infection was confirmed to have emerged in at least 115 roosts across the United States and Canada, spanning over 1,200 km (ref. 11). Bat numbers across affected sites have declined by over 70% and analyses have shown that at least one affected species, the little brown bat *Myotis lucifugus*, has a greater than 99% chance of becoming locally extinct within the next 16 years (ref. 11). Other species of bats across this region are declining as a consequence of this infection, and the prognosis for their survival and their role in supporting healthy ecosystems, is poor¹².

Cases of this sort are no longer perceived to be atypical. The probability of extinction is increasing for some species of North American bats, but another fungal infection has caused the greatest disease-driven loss of biodiversity ever documented. The skin-infecting amphibian fungus

Batrachochytrium dendrobatidis was discovered in 1997 (ref. 13) and named in 1999 (ref. 14). *B. dendrobatidis* has been shown to infect over 500 species of amphibians in 54 countries, on all continents where amphibians are found^{15,16}, and is highly pathogenic across a wide diversity of species. Studies using preserved amphibian specimens showed that the first appearance of *B. dendrobatidis* in the Americas coincided with a wave of population declines that began in southern Mexico in the 1970s and proceeded through Central America to reach the Panamanian isthmus in 2007 (ref. 17). As a consequence of the infection, some areas of central America have lost over 40% of their amphibian species¹⁸, a loss that has resulted in measurable ecosystem-level changes¹⁹. This spatiotemporal pattern has been broadly mirrored in other continents¹⁵, and ongoing reductions in amphibian diversity owing to chytridiomycosis have contributed to nearly half of all amphibian species being in decline worldwide²⁰.





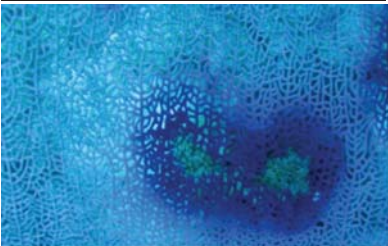


Fungal infections causing widespread population declines are not limited to crops, bats and frogs; studies show that they are emerging as pathogens across diverse taxa (Table 1), including soft corals (for example, sea-fan aspergillosis caused by *Aspergillus sydowii*)²¹, bees (the microsporidian fungus *Nosema* sp. associated with colony collapse disorder)²², and as human and wildlife pathogens in previously non-endemic regions (for example, the emergent virulent VGII lineage of *Cryptococcus gattii* in the northwest America²³ and *Cryptococcus neoformans* across southeast Asia²⁴). The oomycetes have life histories similar to those of fungi and are also emerging as aggressive pathogens of animals, causing declines in freshwater brown crayfish (for example, the crayfish plague caused by *Aphanomyces astaci*)²⁵, Tilapia fish (for example, epizootic ulcerative syndrome caused by *A. invadans*)²⁶ and many species of plants^{27,28}. Although the direct causal relationship is uncertain in some of these diverse host-pathogen relationships, it seems that pathogenic fungi are having a pronounced effect on the global biota¹.

Increasing risk of biodiversity loss by Fungi

For infectious disease systems, theory predicts that pathogens will co-evolve with, rather than extirpate, their hosts^{29,30}. Such evolutionary dynamics mirror population-level processes in which density dependence leads to the loss of pathogens before their hosts are driven extinct³¹. For these reasons, infection has not been widely acknowledged as an

¹Department of Infectious Disease Epidemiology, Imperial College, London W2 1PG, UK. ²Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, California 93106-9620, USA. ³Department of Pediatrics, Harvard Medical School, Children's Hospital Boston, Massachusetts 02115, USA. ⁴ProMED, International Society for Infectious Diseases; and Division of Infectious Diseases and Immunology, University of Massachusetts Medical School, Massachusetts 01655, USA. ⁵Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK.

Table 1 | Major fungal organisms posing threats to animal and plant species.

| | Host | Pathogen (Phylum) | Disease dynamics leading to mass mortality in animal and plant hosts |
|---|---|---|--|
|  | Amphibian species (for example, the common midwife toad, <i>Alytes obstetricans</i>) | <i>Batrachochytrium dendrobatidis</i> (Chytridiomycota) | Worldwide dispersal of a hypervirulent lineage by trade ⁶⁴ . Ultra-generalist pathogen manifesting spillover between tolerant/susceptible species. Extent of chytridiomycosis is dependent on biotic and abiotic context ^{15,82} . |
|  | Rice (<i>Oryza sativa</i>); <i>Magnaporthe grisea</i> species complex on 50 grass and sedge species, including wheat and barley | <i>Magnaporthe oryzae</i> (Ascomycota) | Rice blast disease in 85 countries, causing 10–35% loss of harvest. Global blast population structure determined by deployment of seeds with inbred race-specific disease resistance (RSR). Invasions occur by 'host hops' and altered pathogen demographics. |
|  | Bat spp. (little brown bats, <i>Myotis lucifugus</i>) | <i>Geomyces destructans</i> (Ascomycota) | New invasion of North American bat roosts occurred in approximately 2006, and disease is spreading rapidly ⁸ . Pathogen reservoir may exist in cave soil. Disease is more aggressive compared to similar infections in European bats, possibly owing to differences in roosts and host life histories ⁶⁵ . |
|  | Wheat (<i>Triticum aestivum</i>); 28 <i>Puccinia graminis</i> f. <i>tritici</i> species, but <i>P. graminis</i> is found on 365 cereal or grass species | <i>Puccinia graminis</i> (Basidiomycota) | Wheat stem rust is present on six continents. Population structure is determined by deployment of RSR cultivars and long-distance spread of aeciospores. Strain Ug99 poses a notable threat to resistant wheat varieties, causing up to 100% crop loss. |
|  | Coral species (for example, the sea fan, <i>Gorgonia ventalina</i>) | <i>Aspergillus sydowii</i> (Ascomycota) | Sea-fan aspergillosis caused by a common terrestrial soil fungus ^{21,86} . Epizootics are associated with warm-temperature anomalies. Coral immunosuppression is probably a factor causing decline. |
|  | Bee species (for example, the domestic honeybee (<i>Apis mellifera</i>) suffering colony collapse disorder) | <i>Nosema</i> species (Microsporidia) | Microsporidian fungal infections are associated with colony collapse disorder and declining populations. Pathogen prevalence is probably a part of a multifactorial phenomenon that includes environmental stressors and polyparasitism ^{87,88} . |
|  | Sea turtle species (the loggerhead turtle, <i>Caretta caretta</i>) | <i>Fusarium solani</i> (Ascomycota) | Soil-dwelling saprotroph and phytopathogenic fungus. Infection causes hatch failure in loggerhead turtle nests and suboptimal juveniles ⁴⁴ . The disease dynamics fulfil Koch's postulates. Environmental forcing is suspected but not proven. |

Images in Table 1, with permission: *A. obstetricans* chytridiomycosis mortalities, M.C.F.; *M. oryzae*, N. Talbot; WNS-affected little brown bats, A. Hicks; *P. graminis*, R. Mago; *G. ventalina* infected with *A. sydowii*, D. Harvell; *A. mellifera* hive suffering from colony collapse disorder, J. Evans; sea turtle eggs infected with *F. solani*, J. Diéguez-Urbeondo and A. Marco.

extinction mechanism owing to such intrinsic theoretical biotic limitations³². Inspection of species conservation databases would seem to confirm this idea. The International Union for Conservation of Nature (IUCN) red list database details threats to species worldwide, and analysis of the database has shown that of the 833 recorded species extinctions, less than 4% (31 species) were ascribed to infectious disease⁷. Ecological studies on host–pathogen relationships support this finding by showing that lower parasite richness occurs in threatened host species, suggesting that parasite decline and ‘fade out’ occurs when hosts become rare³³. Therefore, given that macroevolutionary and ecological processes should promote diversity and prevent infectious diseases from driving their host species to extinction, we posed the question of whether we are witnessing increasing disease and extinction events driven by fungi on an increasingly large scale, or, alternatively, if there is evidence that a reporting bias has skewed our opinion of the current level of threat.

EIDs are those pathogens that are increasing in their incidence, geographic or host range, and virulence^{34,35}. Current attempts to detect EID events centre on capturing changes in the patterns of disease alerts recorded by disease monitoring programmes. ProMED (the Program for Monitoring Emerging Diseases; <http://www.promedmail.org>) and HealthMap (<http://healthmap.org>) have two approaches for detecting and monitoring outbreaks worldwide in plant and animal hosts: first, by active reporting of disease outbreaks, and second, by capturing diverse online data sources. To ascertain whether there are changing patterns of fungal disease, we reviewed all disease alerts in ProMED (1994–2010) and HealthMap (2006–10) for combinations of search terms to catalogue fungal alerts. We then classified these according to their relative proportion against the total number of disease alerts, and discriminated between plant- or animal-associated fungal pathogens (Supplementary Table 2). We also searched the primary research literature for reports in which EIDs have caused host extinction events, either at the regional scale (extirpations) or globally (Supplementary Table 3). These analyses show a number of positive trends associated with infectious fungi. Overall, fungal alerts comprise 3.5% of the ~38,000 ProMED records screened. However, over the period from 1995 to 2010, the relative proportion of fungal alerts increased in the ProMED database from 1% to 7% of the database total (Fig. 1a and Supplementary Table 2). This trend is observed for both plant-infecting (0.4% to 5.4%) and animal-infecting (0.5% to 1.4%) fungi. HealthMap shows a recent (2007–11) positive trend in the proportion of records of fungi infecting animals (0.1% to 0.3%) and plants (0.1 to 0.2%), and fungal disease alerts

were shown to occur worldwide (Fig. 1b). Web of Science literature searches and compilation of previous meta-analyses of infection-related species extinction and regional extirpation events show that fungi comprise the highest threat for both animal-host (72%) and plant-host (64%) species (Fig. 1c and Supplementary Tables 3 and 4). This effect is more pronounced for animal hosts (39 animal species affected versus 4 plant species); moreover, there is a notable increase in host loss during the second half of the twentieth century, driven mainly by the emergence of *B. dendrobatidis* (Fig. 1d). This effect is moderated after correcting for mass-species loss in regions of high epizootic loss (such as the mass extirpations of amphibians in Central America). However, fungi remain the major cause (65%) of pathogen-driven host loss after this correction. Our estimates are probably conservative owing to the cryptic nature of most disease-driven species impacts. For example, the lack of disease-related IUCN red list records is due to a lack of baseline data on the incidence of pathogens in natural systems compounded by inadequate disease diagnostics, reporting protocols and a lack of centralized recording mechanisms. Hence, the true numbers of extinctions and extirpations caused by fungi and oomycetes are likely to be greater as we have not been able to categorize the probably high levels of species loss in major plant (such as the *Phytophthora* dieback in Australia caused by *Phytophthora cinnamomi*; Supplementary Table 3) or animal outbreaks (for example, the effects of *B. dendrobatidis* emergence in the American wet tropics). We cannot discount the idea that sampling bias owing to increasing awareness of pathogenic fungi as EIDs may contribute to the patterns that we document. However, because of our observation that increases in the amount of disease caused by fungi are seen across many sources of data, including disease alerts, the peer-reviewed literature and previously noted patterns in human fungal EIDs³⁵, we believe that these trends are real. Therefore, the answer to our question seems to be that the data do indeed support the idea that fungi pose a greater threat to plant and animal biodiversity relative to other taxonomic classes of pathogen and hosts, and that this threat is increasing.

Fungal–disease dynamics leading to host extinction

Here we illustrate several key biological features of fungi that contribute to the epidemiological dynamics underlying contemporary increases in disease emergence and host extinction (Box 1).

High virulence

Fungi, like some bacterial and viral infections, can be highly lethal to naive hosts with rates of mortality approaching 100% (for example,

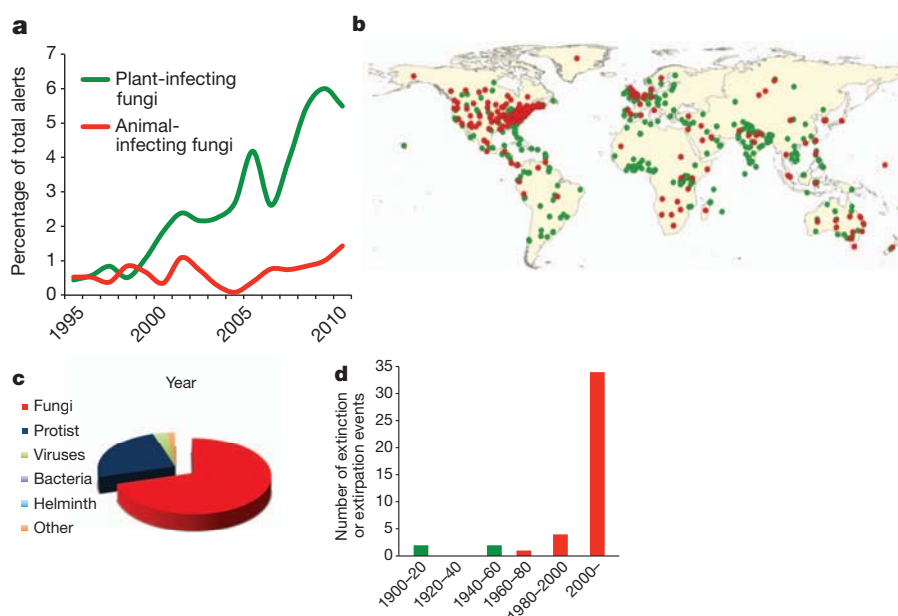


Figure 1 | Worldwide reporting trends in fungal EIDs. **a, b**, Disease alerts in the ProMED database for pathogenic fungi of animals and plants (**a**), and the spatial location of the associated reports (**b**). **c, d**, Relative proportions of species extinction and/or extirpation events for major classes of infectious disease agents (**c**) and their temporal trends for fungal pathogens (**d**). Primary data sources are given in the Supplementary Information.

BOX 1

Modelling host extinctions caused by pathogenic fungi

A simple susceptible–infected model shows that the presence of a threshold host population size for disease persistence does not prevent host extinction during a disease outbreak, especially in cases in which a lethal pathogen invades a large host population. In a large host population transmission is rapid and all hosts can become infected before the host population is suppressed below the threshold. The model follows the dynamics of susceptible (S) and infected (I) hosts during the short time duration of an epidemic (deaths that are not due to disease, and births, are ignored): $dS/dt = -\beta SI$; $dI/dt = \beta SI - \alpha I$, where β is the pathogen transmission rate, and α is the disease-induced death rate. For the parameters shown in Fig. 2a, the threshold population size below which the pathogen has a negative growth rate is $N_T = 20$ individuals. Figure 2a shows that large host populations are rapidly driven extinct, but only a fraction of individuals are killed in small host populations.

Pathogens with a long-lived infectious stage have an increased potential to cause host extinction. In this model, the disease is transmitted through contact between susceptible hosts and free-living infectious spores (Z), resulting in infected hosts, I : $dS/dt = -\beta SZ$; $dI/dt = \beta SZ - \alpha I$; $dZ/dt = \phi I - \mu Z - \beta NZ$, where β is the transmission rate, α is the pathogen-induced death rate and ϕ is the rate of release of spores from infected hosts. Figure 2b shows that fraction of hosts killed in a disease outbreak increases with the duration of the free-living infectious spore stage ($1/\mu$, where μ is the spore mortality rate).

Saprophytic growth by a pathogen can lead to extinction of the host, and even allow the pathogen to persist in the absence of its host. In this model, free-living infectious spores are released from infected hosts (with rate ϕ), and can increase in abundance through saprophytic growth, with rate σ . To illustrate the effects of saprophytic pathogen growth on host and pathogen equilibria (Fig. 2c), density-independent host reproduction (with rate b), density-dependent host mortality (with rate $d_0 + d_1 N$, where $N = S + I$), and density-dependent spore mortalities (at rate $\mu_0 + \mu_1 Z$) were included: $dS/dt = bN - (d_0 + d_1 N)S - \beta SZ$; $dI/dt = \beta SZ - \alpha I - (d_0 + d_1 N)I$; $dZ/dt = \phi I + \sigma Z - (\mu_0 + \mu_1 Z)Z - \beta NZ$.

The presence of a tolerant host species can lead to the extinction of a susceptible host species. In this model, species A is the tolerant host species, which can become infected and shed infectious spores but does not die as a result of the disease, whereas the susceptible host species (species B) has a disease-induced per-capita mortality rate of α_B . Figure 2d shows that species B is driven extinct at high densities of species A. $dS_A/dt = b_A N_A - (d_{A0} + d_{A1} N_A)S_A - \beta_A S_A Z$; $dI_A/dt = \beta_A S_A Z - (d_{A0} + d_{A1} N_A)I_A$; $dS_B/dt = b_B N_B - (d_{B0} + d_{B1} N_B)S_B - \beta_B S_B Z$; $dI_B/dt = \beta_B S_B Z - (d_{B0} + d_{B1} N_B)I_B - \alpha_B I_B$; $dZ/dt = \phi_A I_A + \phi_B I_B - \mu Z - \beta_A N_A Z - \beta_B N_B Z$, where all parameters are as previously defined, but with the subscripts A or B referring to host species A or B, respectively.

B. dendrobatidis in amphibians, *G. destructans* in bats and *Ophiostoma ulmi* in elm trees). Virulence is a measure of the relative capacity of a microbe to cause damage to a host³⁶, and high virulence is associated with rapid intra-host growth rates, ultimately leading to rapid inter-host transmission^{37,38}. Fungi have a high reproductive potential and in a large host population this effect can result in all individuals becoming infected before the population is driven to the low densities at which the pathogen can no longer spread (Fig. 2a). Thus, host extirpation can occur before density dependence limits the rate of transmission, a feature that has contributed to the mass extirpations seen in frog populations across the US Sierra Nevada mountains³⁹. Similarly, even if the pathogen does not drive the host to complete extinction, it may severely reduce the

population size to the point at which the species is vulnerable owing to catastrophic collapses as a result of stochastic³² or Allee effects⁴⁰.

Long-lived environmental stages

Fungi have remarkably resilient dispersal stages (a feature that they share with some spore-forming bacteria, such as *Bacillus anthracis*). The ability to survive independently outside of their host, as a free-living saprophyte or as durable spores in the environment, is probably the most important feature in driving the emergence of pathogenic fungi, owing to an increased risk of transporting the inocula to naive hosts (Fig. 2b)⁴¹. Furthermore, pathogenic fungi with a saprophytic stage (called sapronoses; Fig. 2c) can lead to host extirpation because their growth rate is decoupled from host densities and many fungal diseases threatening natural populations are caused by opportunistic fungi with long-lived environmental stages. Many fungi in the phylum Ascomycota are common soil organisms and are tolerant of salinity with the consequence that, when they enter the marine system through freshwater drainage, they are able to infect susceptible hosts such as corals (*A. sydowii*⁴²), sea otters (*Coccidioides immitis*⁴³) and the nests of loggerhead turtles (*Fusarium solani*⁴⁴). In terrestrial environments, potentially lethal fungi are ubiquitous, such as the causative agent of aspergillosis, *Aspergillus fumigatus*, and soil surveys have shown that *Geomyces* spp. are common soil organisms. Viable *G. destructans* has been recovered from the soil of infected bat caves⁴⁵, showing that the pathogen is able to survive and persist in infected roosts when the bats are absent. Likewise, long-term persistence of fungal inoculum in the agricultural landscape is achieved by quiescent survival on plant debris, such as the spores of wheat stem rust (*Puccinia graminis*), which overwinter on straw stubble before infecting a secondary host.

Generalist pathogens and opportunistic pathogens

Although many fungi demonstrate extreme host specialization, exemplified by the gene-for-gene interactions between biotrophic fungi and their plant hosts, broad host ranges twinned with high virulence can be a lethal combination. Fungi exhibit the broadest spectrum of host ranges for any group of pathogens, and *B. dendrobatidis* and the oomycete *Phytophthora ramorum* (the cause of sudden oak death and ramorum blight) are known to infect 508 (ref. 16) and 109 (ref. 3) host species, respectively. Different host species vary in their susceptibility to infection and these differences create the potential for parasite-mediated competition when the pathogens concerned are generalists⁴⁶. Host species that can tolerate high infection loads while serving as a source of infectious stages (known as pathogen spill-over) act as community 'super spreaders' by maintaining persistent infectious stages in the system (Fig. 2d). Invasive North American signal crayfish, which tolerate infection by the oomycete *A. astaci*, force the infection into more susceptible European species that then decline²⁵, and similarly, although *P. ramorum* is deadly to *Notholithocarpus densiflorus* (tanoak) and several *Quercus* species⁴⁷, many of its other hosts survive infection but generate inoculum themselves for new infections. Furthermore, disease-tolerant life-history stages of otherwise susceptible species can maintain high pathogen levels leading to extinction dynamics. In chytridiomycosis, the long-lived multi-year tadpole stages of amphibians such as the mountain yellow-legged frog *Rana muscosa* and the midwife toad *Alytes obstetricans* are not killed by chytrid infections, but they can build up high loads of *B. dendrobatidis* that can infect and overwhelm juvenile metamorphs of the same species, leading to rapid population loss³⁹. Ultimately, when host-generalist pathogens manifest long-lived environmental stages, conditions may occur that lead to long-distance dispersal and infection of naive hosts and environments⁵.

Trade and transport promotes globalization of fungi

Fungi comprise most of the viable biomass in the air, with an average human breath containing between one and ten fungal spores⁴⁸. This ability of fungi to disperse results in some species with cosmopolitan distributions^{5,49,50}. However, these species are in the minority and it is noticeable that few fungi exhibit truly globally distributions; instead they

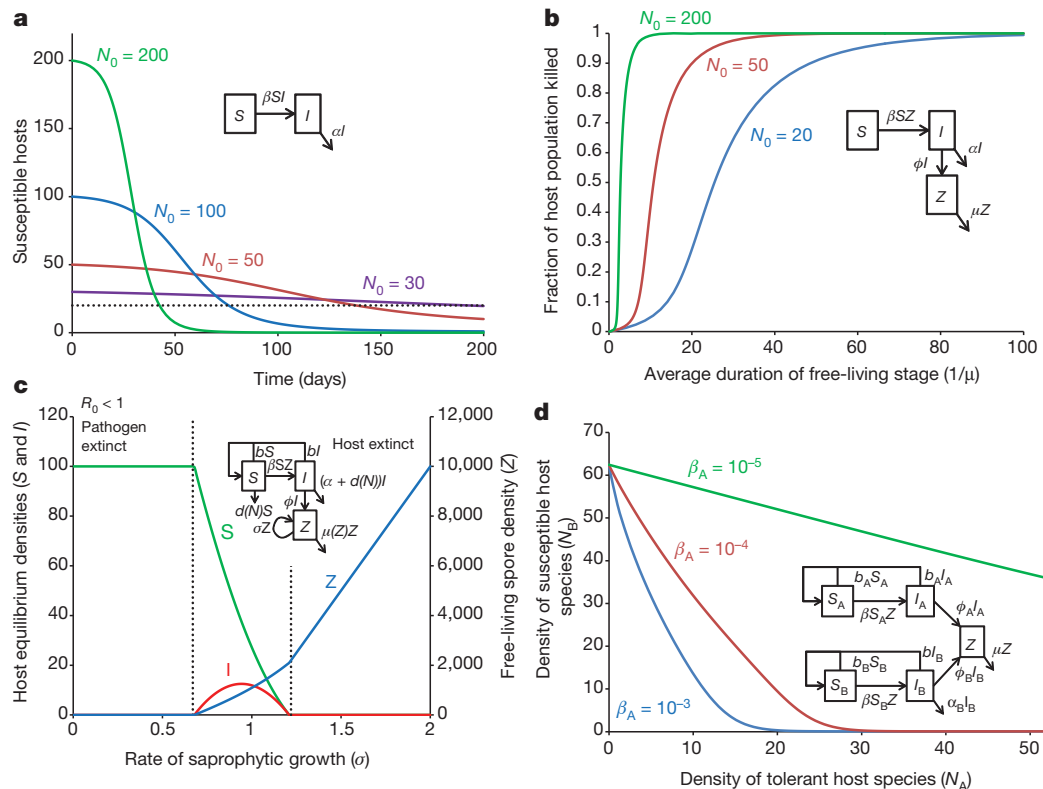


Figure 2 | Fungal disease dynamics leading to host extinction. **a**, The presence of a threshold host population size for disease persistence does not prevent host extinction during a disease outbreak, especially in cases in which a lethal pathogen invades a large host population. In a large host population transmission is rapid and all hosts can become infected before the host population is suppressed below the threshold (pathogen transmission rate, $\beta = 0.001$ per individual per day; disease induced-death rate, $\alpha = 0.02$ per day; simulations start with one infected individual and N_0 susceptible individuals). **b**, Long-lived infectious stages can increase the potential for host extinction. The fraction of hosts killed in a disease outbreak is shown as a function of the duration of the free-living infectious spore stage (pathogen transmission rate, $\beta = 5 \times 10^{-6}$; disease-induced death rate, $\alpha = 0.02$; rate of release of spores from infected hosts, $\phi = 10$; outbreaks initiated with one infected host in a population of N_0 susceptible individuals). **c**, Saprophytic growth: equilibrium densities of susceptible and infected hosts and free-living spores as a function of the rate of saprophytic growth, σ . With no (or low levels of) saprophytic growth, the basic reproductive rate of the pathogen (R_0) is less than 1, the pathogen cannot invade the system and the host persists at its disease-free equilibrium density. Intermediate levels of saprophytic growth allow the pathogen to invade

exhibit spatially restricted endemic ranges⁵¹. In many cases, local adaptation and host specificity are thought to underlie fungal endemism^{51,52}. Nevertheless, when local climatic and vegetative constraints are projected globally it becomes clear that potential ranges of pathogenic fungi may be much larger than their realized range⁵³. If fungi are contained spatially by the combination of physical limits on dispersal, abiotic conditions, host distributions and genetic limits on adaptation, then how are pathogenic fungi able to overcome these barriers? Although fungi have shown the ability to undergo range expansions in response to environmental shifts⁵⁴, human-mediated intercontinental dispersal of unrecognized fungal pathogens is the major component in initiating new chains of transmission.

Pathogenic fungi have dispersed alongside early human migrations, and several thousand years ago two of these fungi, *Coccidioides immitis* and *C. neoformans* lineage VNI, seem to have invaded South America and southeast Asia, respectively, vectored by humans and their domesticated animals^{24,55}. Similar ancient patterns of human-associated disease spread are detected by studies of the genome diversity of many plant fungal pathogens⁵⁶. However, more recent increases in fungal disease are

attributable to the many-fold increase in fungal-infected trade products and food⁵⁷. The consequences of recent introductions of pathogens in association with trade are well known; examples include the Irish Famine⁴ (a consequence of *Phytophthora infestans* late blight introduction from South America), the destruction of the North American chestnuts⁵⁸ (caused by the importation of *Cryphonectria parasitica*-infected Asian chestnut trees to the east coast of the United States in the early twentieth century) and the Second World War introduction of *Heterobasidion annosum* into Italy from the USA (vectored by untreated wooden transport crates)⁵⁹. Human-mediated intercontinental trade has also been linked clearly to the spread of animal-pathogenic fungi through the transportation of infected vector species. *B. dendrobatidis* has been introduced repeatedly to naive populations worldwide as a consequence of the trade in the infected, yet disease-tolerant species such as North American bullfrogs (*Rana catesbeiana*)^{60–62} and African clawed frogs (*Xenopus laevis*)^{63,64}. Whether the emergence of bat WNS constitutes an introduction of *G. destructans* into North America from Europe or elsewhere remains to be shown. However, the widespread but apparently non-pathogenic nature of the infection in European bats tentatively

suggests that the disease may have been vectored from this region in contaminated soil⁶⁵.

Accelerated evolution of virulence in pathogenic fungi

Human activities are not only associated with the dispersal of pathogenic fungi, they also interact with key fungal characteristics, such as habitat flexibility, environmental persistence and multiple reproductive modes, to cause the emergence of disease. Importantly, many fungi are flexible in their ability to undergo genetic recombination, hybridization or horizontal gene transfer⁶⁶, causing the clonal emergence of pathogenic lineages but also allowing the formation of novel genetic diversity leading to the genesis of new pathogens^{56, 67}. Reproductive barriers in fungi are known to evolve more rapidly between sympatric lineages that are in the nascent stages of divergence than between geographically separated allopatric lineages, in a process known as reinforcement^{68, 69}. As a consequence, anthropogenic mixing of previously allopatric fungal lineages that still retain the potential for genetic exchange can drive rapid macroevolutionary change. Although many hybrids are inviable owing to genome incompatibilities, large phenotypic leaps can be achieved by the resulting 'hopeful monsters', leading to host jumps and increased virulence⁷⁰. Such mechanisms are thought to drive the formation of new pathotypes in plant pathogens⁵², and oomycetes as well as fungi exhibit the genesis of new interspecific hybrids as lineages come into contact^{71, 72}. Evidence of the effect of multiple fungal co-dispersal events and recombination can also be seen in the recent *C. gattii* outbreaks in northwestern North America. In this case, strains that do not normally recombine have increased their virulence by undergoing recombination and adaptation to overcome mammalian immune responses^{23, 67}. Recent studies based on the resequencing of *B. dendrobatidis* genomes have shown that, although several lineages exist, only a single lineage (known as the *B. dendrobatidis* global panzootic lineage) has emerged in at least five continents during the twentieth century to cause epizootic amphibian declines⁶⁴. Notably, the genome of the *B. dendrobatidis* global panzootic lineage shows the hallmarks of a single hybrid origin and, when compared against other newly discovered lineages of *B. dendrobatidis*, is more pathogenic, suggesting that transmission and onward spread of the lineage has been facilitated by an increase in its virulence. Given that the rate of intra- and inter-lineage recombination among fungi will be proportional to the contact rates between previously geographically separate populations and species, these data from across plant and animal fungal pathosystems suggest that the further evolution of new races is inevitable given current rates of homogenization of previously allopatric, geographically separated, fungal lineages.

Environmental change as a driver of fungal EIDs

Climate fluctuation can be a potent cofactor in forcing changing patterns of fungal phenology⁷³ and are known to govern plant fungal EIDs. Models of climate change for the coming decades predict increases in global temperature, atmospheric CO₂, ozone and changes in humidity, rainfall and severe weather⁷⁴. For this reason, many interactions must be taken into consideration when attempting to predict the future effects of climate change on plant diseases⁷⁵. First, the physiological and spatial changes that plants may undergo in response to the various different components of climate change and the resulting effects on the pathogen⁷⁶, and second, the effects on the pathogen's physiology and dispersal external to their plant hosts⁷⁵. Frequently, however, experimental models have only taken into account one element of climate change, a common example being the free-air CO₂ enrichment (FACE) studies that model the effects of elevated atmospheric CO₂ (ref. 77). A notable result here has been rice blast severity being higher at higher CO₂ levels⁷⁸. However, although there has been a general trend for increased disease severity under simulated climate-change conditions⁷⁹, and although some species are thought to be changing their distribution in response to these changes (for example, *P. graminis*⁸⁰), other elements of climate change, such as

increased ozone, have been shown to have the opposite effect (for example, in *Puccinia recondita*⁷⁷).

Evidence for the idea that climate change has an impact on the dynamics and distribution of animal-infecting fungi is less clear-cut than that in relation to plant-infecting fungi and, although arguments have been made that warming trends may have contributed to the emergence of *B. dendrobatidis* in Central America and Europe^{81, 82}, there is active debate about these conclusions^{83, 84}. Regardless, it is clear that the disease state, chytridiomycosis, is linked to environmental factors; regional climate warming can increase the local range of the pathogen⁵⁴ and disease risk is inversely related to rates of deforestation⁸⁵. Correlations between ecosystem change and a rise in infection by opportunistic pathogens has been proposed to account for the occurrence of coral reef declines worldwide. For example, disease caused by a variety of microbes threatens hard corals to the extent that two of the most abundant Caribbean reef-builders (staghorn and elkhorn corals) are now listed under the US Endangered Species Act. Across varied reef systems, the occurrence of warm-temperature anomalies leading to bleaching events is associated with increases in disease caused by opportunistic pathogens such as *A. sydowii*⁸⁶. In an allied colonial system, colony collapse disorder has resulted in steep declines of the European honeybee *Apis mellifera* in Europe and North America⁸⁷. These losses seem to be influenced by a mixture of aetiological agents that are fungal (for example, microsporidian (*Nosema ceranae*)), viral (for example, Kashmir bee virus and Israeli acute paralysis virus) and ectoparasitic (for example, *Varroa destructor*) in origin. So far, no single environmental cause has been identified that can account for the apparent reduction in the ability of honeybee colonies to resist these infections, and agricultural chemicals, malnutrition and modern beekeeping practices have all been suggested as potential cofactors for colony-collapse disorder⁸⁸. The increasing use of azole-based agricultural chemicals has been implicated as a factor underpinning the increase in the frequency of multiple-triazole-resistant (MTR) isolates of *A. fumigatus* infecting humans⁸⁹. The widespread agricultural use of azoles as a means of combating crop pathogens is speculated to have led to selection for MTR alleles, an idea that is supported by the recent discovery that resistance clusters onto a single lineage in Dutch populations of the fungus⁹⁰. Efforts must now be turned to integrating epidemiological studies with those on environmental change so that the many possible interactions and outcomes can be assessed, as making blanket predictions for fungal diseases is currently impossible⁹¹. The highly coordinated response to the recent outbreak of wheat stem rust (*P. graminis*, strain Ug99) is a positive step towards this goal^{77, 92}.

Fungal EIDs impact food security and ecosystem services

Impacts of fungal diseases are clearly manifested in crops and there are direct measurable economic consequences associated with die-off in forest and urban environments. Losses that are due to persistent and epidemic outbreaks of fungal and oomycete infection in rice (rice blast caused by *Magnaporthe oryzae*), wheat (rust caused by *P. graminis*), maize (smut caused by *Ustilago maydis*), potatoes (late blight caused by *P. infestans*) and soybean (rust caused by *Phakospora pachyrizii*) vary regionally but pose a current and growing threat to food security². Our estimates of loss of food are based on the 2009–10 world harvest statistics of five of our most important crops and make certain basic assumptions of calorific value and worldwide average production (Supplementary Table 1). Our calculations show that even low-level persistent disease leads to losses that, if mitigated, would be sufficient to feed 8.5% of the 7 billion humans alive in 2011. If severe epidemics in all five crops were to occur simultaneously, this would leave food sufficient for only 39% of the world's population, but the probability of such an event occurring is very low indeed.

Invasive tree diseases have caused the loss of approximately 100 million elm trees in the United Kingdom and the United States^{52, 93}, and 3.5 billion chestnut trees have succumbed to chestnut blight in the United States (Supplementary Table 5). Losses of western Canadian pine trees to the mountain pine beetle–blue-stain fungus association will result in the

release of 270 megatonnes of CO₂ over the period from 2000 to 2020, with a clearly ascribed economic cost both for the wood itself and the carbon released⁹⁴. These, and other diseases such as 'sudden oak death' in California and 'foliar and twig blight' and 'dieback' on ornamental trees, woody shrubs and forestry plants in the European Union, affect ecological diversity, are costly to manage and account for huge losses of fixed CO₂. Indeed, we calculate regional losses of absorbed CO₂ to total 230–580 megatonnes for just a handful of diseases (Supplementary Table 5) with the higher figure equating to 0.069% of the global atmospheric CO₂. We have included both emerging (Jarrah dieback, sudden oak death and pine beetle–blue-stain fungus) and emergent diseases (Dutch elm blight and chestnut blight), as these represent the few examples for which informed estimates are possible. We are unable to quantify any of the many other recent emerging diseases, such as red band needle blight of pines, *Phytophthora alni* on alders or pitch pine canker on Monterey pines, owing to a lack of data and economic interest, both of which are trends that must be reversed. Assessing the economic burden of fungal mycoses in animals is a challenging task. Although the impact of fungal EIDs is manifested in domestic animal settings, particularly the amphibian trade⁹⁵ and in regions where virulent lineages have established⁹⁶, reporting mechanisms for outbreaks do not widely exist. In natural settings, valuations have recently estimated the losses to US agriculture that are the result of declines in bat populations at more than US\$3.7 billion per year (ref. 12). However, although broad ecosystem-level impacts of other fungal EIDs of wildlife are suspected, economic valuations of the ecosystem services that these species support are wholly lacking.

Mitigating fungal EIDs in animals and plants

The high socioeconomic value of crops means that detection and control of fungal diseases in agriculture far outpaces that in natural habitats. Epidemiological models have been developed to predict the risk of seasonally specific crop pathogens, allowing targeted control, and specific threats are assessed through consortia of research, governmental and global non-governmental organizations, led by the United Nations Food and Agricultural Organization (FAO), and related organizations. Scientifically led development of disease-resistant crop varieties has been mainly successful, although monocultures have in some instances vastly increased the susceptibility of harvests to highly virulent pathogens, a pertinent example being *P. graminis* Ug99. Conversely, although there have been some attempts to mitigate the fungal disease burden in wildlife *in situ*—most notably efforts to eliminate *B. dendrobatidis* in infected populations with the antifungal itraconazole⁹⁷ and the use of probiotic bacteria⁹⁸—communicable wildlife EIDs are essentially unstoppable once they have emerged. International biosecurity against the spread of plant fungal pathogens, although not perfect, is more advanced than protocols to protect against the introduction of animal-associated fungi. Fundamentally, this is the result of a financial dynamic: wildlife is not currently valued economically, whereas crops are.

The World Organisation for Animal Health (also known as the OIE) and the FAO may be the best-placed authorities to coordinate tighter biosecurity controls for trade-associated fungal pathogens of animals. The OIE has listed *B. dendrobatidis* and the crayfish pathogen *A. astaci* in the Aquatic Animal Health Code as internationally notifiable infections, and the FAO compiles outbreak data on transboundary animal diseases using the emergency prevention information system (EMPRES-i). Similarly, the IUCN Wildlife Health Specialist Group determines policy that is specific to combating emerging wildlife disease internationally. On national scales there are a number of initiatives being deployed and in the United States the National Wildlife Health Centre has developed the national federal plan⁹⁹ to mitigate WNS in bats. Intensive monitoring and surveillance will be increasingly important in the coming years because predictive modelling and small-scale experiments can never fully predict future disease spread and severity. An increased political and public profile for the effects of fungal diseases in natural habitats is needed to highlight the importance of fungal disease control outside of the managed agricultural environment to

policy makers. If this occurs, then there will be more sympathy for attempts to improve the regulatory frameworks that are associated with biosecurity in international trade, as this is the most important tool to tackle both plant and animal fungal EIDs now and in the future. The monitoring of fungal inocula in wild populations should be the utmost priority and tighter control of international trade in biological material must be imposed, and with considerable haste. Inadequate biosecurity will mean that new fungal EIDs and virulent races will emerge at an increasingly destructive rate. In addition to better global monitoring and control, attention must also be turned to increasing our understanding of the interactions between hosts, pathogens and the environment, across regional and global scales. Integrated approaches encompassing theoretical and practical epidemiology, climate forecasting, genomic surveillance and monitoring molecular evolution are needed. These should be facilitated by scientists from currently disparate research fields entering into regular global discussions to develop clear and urgent strategies for working towards the elusive magic bullet for emerging fungal diseases: effective prevention and timely control.

1. The Institute of Medicine. *Fungal Diseases: an Emerging Threat to Human Animal and Wildlife Health* (National Academy of Sciences, 2011).
2. Pennisi, E. Armed and dangerous. *Science* **327**, 804–805 (2010).
3. Grünwald, N. J., Goss, E. M. & Press, C. M. *Phytophthora ramorum*: a pathogen with a remarkably wide host range causing sudden oak death on oaks and ramorum blight on woody ornamentals. *Mol. Plant Pathol.* **9**, 729–740 (2008).
4. Anderson, P. K. *et al.* Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol. Evol.* **19**, 535–544 (2004).
5. Brown, J. K. M. & Hovmoller, M. S. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297**, 537–541 (2002).
6. Daszak, P., Cunningham, A. A. & Hyatt, A. D. Emerging infectious diseases of wildlife—threats to biodiversity and human health. *Science* **287**, 443–449 (2000).
7. Smith, K. F., Sax, D. F. & Lafferty, K. D. Evidence for the role of infectious disease in species extinction and endangerment. *Conserv. Biol.* **20**, 1349–1357 (2006).
8. Blehert, D. S. *et al.* Bat white-nose syndrome: an emerging fungal pathogen? *Science* **323**, 227 (2009).
9. Gargas, A., Trest, M. T., Christensen, M., Volk, T. J. & Blehert, D. S. *Geomyces destructans* sp. nov. associated with bat white-nose syndrome. *Mycotaxon* **108**, 147–154 (2009).
10. Lorch, J. M. *et al.* Experimental infection of bats with *Geomyces destructans* causes white-nose syndrome. *Nature* **480**, 376–378 (2011).
11. Frick, W. F. *et al.* An emerging disease causes regional population collapse of a common North American bat species. *Science* **329**, 679–682 (2010).
12. Boyles, J. G., Cryan, P. M., McCracken, G. F. & Kunz, T. H. Economic importance of bats in agriculture. *Science* **332**, 41–42 (2011).
13. Berger, L. *et al.* Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proc. Natl Acad. Sci. USA* **95**, 9031–9036 (1998).
14. Longcore, J. E., Pessier, A. P. & Nichols, D. K. *Batrachochytrium dendrobatidis* gen. et sp. Nov., a chytrid pathogenic to amphibians. *Mycologia* **91**, 219–227 (1999).
15. Fisher, M. C., Garner, T. W. J. & Walker, S. F. Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. *Annu. Rev. Microbiol.* **63**, 291–310 (2009).
16. Bd-Maps. (<http://www.bd-maps.net/>) (accessed, February 2012).
17. Cheng, T. L., Rovito, S. M., Wake, D. B. & Vredenburg, V. T. Coincident mass extirpation of neotropical amphibians with the emergence of the infectious fungal pathogen *Batrachochytrium dendrobatidis*. *Proc. Natl Acad. Sci. USA* **108**, 9502–9507 (2011).
18. Crawford, A. J., Lips, K. R. & Bermingham, E. Epidemic disease decimates amphibian abundance, species diversity, and evolutionary history in the highlands of central Panama. *Proc. Natl Acad. Sci. USA* **107**, 13777–13782 (2010).
19. Colón-Gaud, C. *et al.* Assessing ecological responses to catastrophic amphibian declines: patterns of macroinvertebrate production and food web structure in upland Panamanian streams. *Limnol. Oceanogr.* **54**, 331–343 (2009).
20. Stuart, S. N. *et al.* Status and trends of amphibian declines and extinctions worldwide. *Science* **306**, 1783–1786 (2004).
21. Kim, K. & Harvell, C. D. The rise and fall of a six-year coral-fungal epizootic. *Am. Nat.* **164**, S52–S63 (2004).
22. Cameron, S. A. *et al.* Patterns of widespread decline in North American bumble bees. *Proc. Natl Acad. Sci. USA* **108**, 662–667 (2011).
23. Byrnes, E. J. III *et al.* Emergence and pathogenicity of highly virulent *Cryptococcus gattii* genotypes in the northwest United States. *PLoS Pathog.* **6**, e1000850 (2010).

24. Simwami, S. P. *et al.* Low diversity *Cryptococcus neoformans* variety *grubii* multilocus sequence types from Thailand are consistent with an ancestral African origin. *PLoS Pathog.* **7**, e1001343 (2011).
25. Holdich, D. M., Reynolds, J. D., Souty-Grosset, C. & Sibley, P. J. A review of the ever increasing threat to European crayfish from non-indigenous crayfish species. *Knowl. Managt. Aquat. Ecosyst.* **394–395**, 11 (2009).
26. Andrew, T. G., Huchzermeyer, K. D. A., Mbeha, B. C. & Nengu, S. M. Epizootic ulcerative syndrome affecting fish in the Zambezi river system in southern Africa. *Vet. Rec.* **163**, 629–631 (2008).
27. Rizzo, D. M. & Garbelotto, M. Sudden oak death: endangering California and Oregon forest ecosystems. *Front. Ecol. Environ.* **1**, 197–204 (2003).
28. Wills, R. T. The ecological impact of *Phytophthora cinnamomi* in the Stirling Range National Park, Western Australia. *Aust. J. Ecol.* **18**, 145–159 (1993).
29. Jaenike, J. An hypothesis to account for the maintenance of sex within populations. *Evol. Theor.* **3**, 191–194 (1978).
30. Paterson, S. *et al.* Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278 (2010).
31. McCallum, H. & Dobson, A. Detecting disease and parasite threats to endangered species and ecosystems. *Trends Ecol. Evol.* **10**, 190–194 (1995).
32. De Castro, F. & Bolker, B. Mechanisms of disease-induced extinction. *Ecol. Lett.* **8**, 117–126 (2005).
- Theoretical study outlining the conditions under which disease can cause extinction of its host species.**
33. Altizer, S., Nunn, C. L. & Lindenberg, P. Do threatened hosts have fewer parasites? A comparative study in primates. *J. Anim. Ecol.* **76**, 304–314 (2007).
34. Daszak, P., Cunningham, A. A. & Hyatt, A. D. Emerging infectious diseases of wildlife—threats to biodiversity and human health. *Science* **287**, 443–449 (2000).
35. Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
- Macroecological analysis of recent patterns of EIDs worldwide in humans.**
36. Casadevall, A. & Pirofski, L. A. The damage response framework of microbial pathogenesis. *Nature Rev. Microbiol.* **1**, 17–24 (2003).
37. de Roode, J. C. *et al.* Virulence and competitive ability in genetically diverse malaria infections. *Proc. Natl Acad. Sci. USA* **102**, 7624–7628 (2005).
38. Nowak, M. A. & May, R. M. Superinfection and the evolution of parasite virulence. *Proc. R. Soc. Lond. B* **255**, 81–89 (1994).
39. Briggs, C. J., Knapp, R. A. & Vredenburg, V. T. Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. *Proc. Natl Acad. Sci. USA* **107**, 9695–9700 (2010).
40. Stephens, P. A., Sutherland, W. J. & Freckleton, R. P. What is the Allee effect? *Oikos* **87**, 185–190 (1999).
41. Mitchell, K. M., Churcher, T. S., Garner, T. W. G. & Fisher, M. C. Persistence of the emerging pathogen *Batrachochytrium dendrobatidis* outside the amphibian host greatly increases the probability of host extinction. *Proc. R. Soc. B* **275**, 329–334 (2008).
42. Rypien, K. L., Andras, J. P. & Harvell, C. D. Globally panmictic population structure in the opportunistic fungal pathogen *Aspergillus sydowii*. *Mol. Ecol.* **17**, 4068–4078 (2008).
43. Jessup, D. A. *et al.* Southern sea otter as a sentinel of marine ecosystem health. *EcoHealth* **1**, 239–245 (2004).
44. Sarmiento-Ramírez, J. M. *et al.* *Fusarium solani* is responsible for mass mortalities in nests of loggerhead sea turtle, *Caretta caretta*, in Boavista, Cape Verde. *FEMS Microbiol. Lett.* **312**, 192–200 (2010).
45. Lindner, D. L. *et al.* DNA-based detection of the fungal pathogen *Geomyces destructans* in soils from bat hibernacula. *Mycologia* **103**, 241–246 (2011).
46. Holt, R. D., Dobson, A. P., Begon, M., Bowers, R. G. & Schaub, E. M. Parasite establishment in host communities. *Ecol. Lett.* **6**, 837–842 (2003).
47. Hansen, E. M., Parke, J. L. & Sutton, W. Susceptibility of Oregon forest trees and shrubs to *Phytophthora ramorum*: a comparison of artificial inoculation and natural infection. *Plant Dis.* **89**, 63–70 (2005).
48. Fröhlich-Nowoisky, J., Pickersgill, D. A., Despres, V. R. & Poschl, U. High diversity of fungi in air particulate matter. *Proc. Natl Acad. Sci. USA* **106**, 12814–12819 (2009).
49. Henk, D. A. *et al.* Speciation despite globally overlapping distributions in *Penicillium chrysogenum*: the population genetics of Alexander Fleming's lucky fungus. *Mol. Ecol.* **20**, 4288–4301 (2011).
50. Pringle, A., Baker, D. M., Platt, J. L., Latge, J. P. & Taylor, J. W. Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. *Evolution* **59**, 1886–1899 (2005).
51. Ellison, C. E. *et al.* Population genomics and local adaptation in wild isolates of a model microbial eukaryote. *Proc. Natl Acad. Sci. USA* **108**, 2831–2836 (2011).
52. Giraud, T., Gladieux, P. & Gavrilov, S. Linking the emergence of fungal plant diseases with ecological speciation. *Trends Ecol. Evol.* **25**, 387–395 (2010).
53. Springer, D. J. & Chaturvedi, V. Projecting global occurrence of *Cryptococcus gattii*. *Emerg. Infect. Dis.* **16**, 14–20 (2010).
54. Seimon, T. A. *et al.* Upward range extension of Andean anurans and chytridiomycosis to extreme elevations in response to tropical deglaciation. *Glob. Change Biol.* **13**, 288–299 (2007).
55. Fisher, M. C. *et al.* Biogeographic range expansion into South America by *Coccidioides immitis* mirrors New World patterns of human migration. *Proc. Natl Acad. Sci. USA* **98**, 4558–4562 (2001).
56. Stukenbrock, E. H. & McDonald, B. A. The origins of plant pathogens in agro-ecosystems. *Annu. Rev. Phytopathol.* **46**, 75–100 (2008).
57. Brasier, C. M. The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathol.* **57**, 792–808 (2008).
- An analysis of the lack of biosecurity and of the risk of disease introduction associated with the international plant trade.**
58. Milgroom, M. G., Wang, K. R., Zhou, Y., Lipari, S. E. & Kaneko, S. Intercontinental population structure of the chestnut blight fungus, *Cryphonectria parasitica*. *Mycologia* **88**, 179–190 (1996).
59. Gonthier, P., Warner, R., Nicolotti, G., Mazzaglia, A. & Garbelotto, M. M. Pathogen introduction, as a collateral effect of military activity. *Mycol. Res.* **108**, 468–470 (2004).
60. Goka, K. *et al.* Amphibian chytridiomycosis in Japan: distribution, haplotypes and possible route of entry into Japan. *Mol. Ecol.* **18**, 4757–4774 (2009).
61. Garner, T. W. J. *et al.* The emerging amphibian pathogen *Batrachochytrium dendrobatidis* globally infects introduced populations of the North American bullfrog, *Rana catesbeiana*. *Biol. Lett.* **2**, 455–459 (2006).
62. Cunningham, A. A. *et al.* Emergence of amphibian chytridiomycosis in Britain. *Vet. Rec.* **157**, 386–387 (2005).
63. Walker, S. F. *et al.* Invasive pathogens threaten species recovery programs. *Curr. Biol.* **18**, R853–R854 (2008).
64. Farrer, R. A. *et al.* Multiple emergences of genetically diverse amphibian-infecting chytrids include a globalised hypervirulent recombinant lineage. *Proc. Natl Acad. Sci. USA* **108**, 18732–18736 (2011).
- Population genomics analysis of the generation, and spread, of a hypervirulent fungal lineage in amphibians worldwide.**
65. Wibbelt, G. *et al.* White-nose syndrome Fungus (*Geomyces destructans*) in Bats, Europe. *Emerg. Infect. Dis.* **16**, 1237–1243 (2010).
66. Richards, T. A. *et al.* Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proc. Natl Acad. Sci. USA* **108**, 15258–15263 (2011).
67. Fraser, J. A. *et al.* Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak. *Nature* **437**, 1360–1364 (2009).
- Analysis of the evolution of a hypervirulent lineage of human-infecting fungus that invaded British Columbia.**
68. Turner, E., Jacobson, D. J. & Taylor, J. W. Genetic architecture of a reinforced, postmating, reproductive isolation barrier between *Neurospora* species indicates evolution via natural selection. *PLoS Genet.* **7**, e1002204 (2011).
69. Coyne, J. A. & Orr, H. A. *Speciation* (Sinauer Associates, 2004).
70. Mallet, J. Hybrid speciation. *Nature* **446**, 279–283 (2007).
71. Brasier, C. M., Rose, J. & Gibbs, J. N. An unusual phytophthora associated with widespread alder mortality in Britain. *Plant Pathol.* **44**, 999–1007 (1995).
72. Inderbitzin, P., Davis, R. M., Bostock, R. M. & Subbarao, K. V. The ascomycete *Verticillium longisporum* is a hybrid and a plant pathogen with an expanded host range. *PLoS One* **6**, e18260 (2011).
73. Gange, A. C., Gange, E. G., Sparks, T. H. & Boddy, L. Rapid and recent changes in fungal fruiting patterns. *Science* **316**, 71 (2007).
74. Pachauri, R. K. & Resinger, A. (eds) Climate change 2007: Synthesis report. (Intergovernmental Panel on Climate Change, 2007).
75. Newton, A. C., Johnson, S. N. & Gregory, P. J. Implications of climate change for diseases, crop yields and food security. *Euphytica* **179**, 3–18 (2011).
- This paper highlights the importance of understanding the impact of climate change on crops and disease.**
76. Lake, J. A. & Wade, R. N. Plant–pathogen interactions and elevated CO₂: morphological changes in favour of pathogens. *J. Exp. Bot.* **60**, 3123–3131 (2009).
77. Chakraborty, S. *et al.* Impacts of global change on diseases of agricultural crops and forest trees. *CAB Rev.* **3**, 1–5 (2008).
78. Kobayashi, T. *et al.* Effects of elevated atmospheric CO₂ concentration on the infection of rice blast and sheath blight. *Phytopathology* **96**, 425–431 (2006).
79. Madgwick, J. W. *et al.* Impacts of climate change on wheat anthesis and fusarium ear blight in the UK. *Eur. J. Plant Pathol.* **130**, 117–131 (2011).
80. Gregory, P. J., Johnson, S. N., Newton, A. C. & Ingram, J. S. I. Integrating pests and pathogens into the climate change/food security debate. *J. Exp. Bot.* **60**, 2827–2838 (2009).
81. Pounds, J. A. *et al.* Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* **439**, 161–167 (2006).
82. Bosch, J., Carrascal, L. M., Duran, L., Walker, S. & Fisher, M. C. Climate change and outbreaks of amphibian chytridiomycosis in a montane area of Central Spain; is there a link? *Proc. R. Soc. B* **274**, 253–260 (2007).
83. Rohr, J. R., Raffel, T. R., Romansic, J. M., McCallum, H. & Hudson, P. J. Evaluating the links between climate, disease spread, and amphibian declines. *Proc. Natl Acad. Sci. USA* **105**, 17436–17441 (2008).
84. Garner, T. W. J., Rowcliffe, J. M. & Fisher, M. C. Climate change, chytridiomycosis or condition: an experimental test of amphibian survival. *Glob. Change Biol.* **17**, 667–675 (2011).
85. Becker, C. G. & Zamudio, K. R. Tropical amphibian populations experience higher disease risk in natural habitats. *Proc. Natl Acad. Sci. USA* **108**, 9893–9898 (2011).
86. Harvell, C. D. *et al.* Review: Emerging marine diseases - Climate links and anthropogenic factors. *Science* **285**, 1505–1510 (1999).
87. vanEngelsdorp, D. *et al.* Colony collapse disorder: a descriptive study. *PLoS One* **4**, e6481 (2009).
88. Ratnieks, F. L. W. & Carreck, N. L. Clarity on honey bee collapse? *Science* **327**, 152–153 (2010).
89. Verweij, P. E., Mellado, E. & Melchers, W. J. G. Multiple-triazole-resistant aspergillosis. *N. Engl. J. Med.* **356**, 1481–1483 (2007).
90. Klaassen, C. H. W., Gibbons, J. G., Fedorova, N. D., Meis, J. F. & Rokas, A. Evidence for genetic differentiation and variable recombination rates among Dutch populations of the opportunistic human pathogen *Aspergillus fumigatus*. *Mol. Ecol.* **21**, 57–70 (2012).
91. Miraglia, M. *et al.* Climate change and food safety: an emerging issue with special focus on Europe. *Food Chem. Toxicol.* **47**, 1009–1021 (2009).
92. Stokstad, E. The famine fighter's last battle. *Science* **324**, 710–712 (2009).

93. Loo, J. A. Ecological impacts of non-indigenous invasive fungi as forest pathogens. *Biol. Invasions* **11**, 81–96 (2009).
94. Kurz, W. A. *et al.* Mountain pine beetle and forest carbon feedback to climate change. *Nature* **452**, 987–990 (2008).
- Study describing pest- and pathogen-induced loss of forest carbon sinks.**
95. Mazzoni, R. *et al.* Emerging pathogen of wild amphibians in frogs (*Rana catesbeiana*) farmed for international trade. *Emerg. Infect. Dis.* **9**, 995–998 (2003).
96. Byrnes, E. J. III, Bildfell, R. J., Dearing, P. L., Valentine, B. A. & Heitman, J. *Cryptococcus gattii* with bimorphic colony types in a dog in western Oregon: additional evidence for expansion of the Vancouver Island outbreak. *J. Vet. Diagn. Invest.* **21**, 133–136 (2009).
97. Lubick, N. Emergency medicine for frogs. *Nature* **465**, 680–681 (2010).
98. Harris, R. N. *et al.* Skin microbes on frogs prevent morbidity and mortality caused by a lethal skin fungus. *ISME J.* **3**, 818–824 (2009).
99. U.S. Fish and Wildlife Service. A national plan for assisting states, federal agencies, and tribes in managing white-nose syndrome in bats. (<http://www.fws.gov/WhiteNoseSyndrome/pdf/WNSnationalplanMay2011.pdf>) (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements M.C.F. was supported by grants from the Wellcome Trust, Natural Environment Research Council (NERC), and the European Research Area (ERA)-net project BiodivERsA. D.A.H. was supported by a grant from the Leverhulme Trust, C.J.B. was supported by the US National Science Foundation (NSF) Ecology of Infectious Disease grant EF-0723563. S.J.G. was supported by grants from the UK Biotechnology and Biological Sciences Research Council (BBSRC) and the John Fell Fund of the University of Oxford, and S.L.M. was supported by a graduate scholarship from Magdalen College, University of Oxford. J.S.B. was supported by Google.org and the National Institutes of Health grant 5R01LM010812-02. N. Knowlton and J. Heitman provided impetus to develop this review.

Author Contributions M.C. F., D.A.H., C.J.B., S.L.M. and S.J.G. designed, analysed and wrote the paper. Data were collected and analysed by J.S.B. and L.C.M.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to M.C.F. (matthew.fisher@imperial.ac.uk) or S.J.G. (sarah.gurr@plants.ox.ac.uk).

An elementary quantum network of single atoms in optical cavities

Stephan Ritter¹, Christian Nölleke¹, Carolin Hahn¹, Andreas Reiserer¹, Andreas Neuzner¹, Manuel Uphoff¹, Martin Mücke¹, Eden Figueroa¹, Joerg Bochmann^{1†} & Gerhard Rempe¹

Quantum networks are distributed quantum many-body systems with tailored topology and controlled information exchange. They are the backbone of distributed quantum computing architectures and quantum communication. Here we present a prototype of such a quantum network based on single atoms embedded in optical cavities. We show that atom-cavity systems form universal nodes capable of sending, receiving, storing and releasing photonic quantum information. Quantum connectivity between nodes is achieved in the conceptually most fundamental way—by the coherent exchange of a single photon. We demonstrate the faithful transfer of an atomic quantum state and the creation of entanglement between two identical nodes in separate laboratories. The non-local state that is created is manipulated by local quantum bit (qubit) rotation. This efficient cavity-based approach to quantum networking is particularly promising because it offers a clear perspective for scalability, thus paving the way towards large-scale quantum networks and their applications.

Connecting individual quantum systems via quantum channels creates a quantum network with properties profoundly different from any classical network. First, the accessible state space increases exponentially with the number of constituents. Second, the distribution of quantum states across the whole network leads to non-local correlations. Further, the quantum channels mediate long-range or even infinite-range interactions, which can be switched on and off at will. This makes quantum networks tailor-made quantum many-body systems with adjustable degrees of connectivity and arbitrary topologies, and thus powerful quantum simulators. Open questions like the scaling behaviour, percolation of entanglement¹, multi-partite entanglement^{2,3} and quantum phase transitions^{4–6} make quantum networks a prime theme of current theoretical and experimental research. Similarly, quantum networks form the basis of quantum communication and distributed quantum information processing architectures, with interactions taking the form of quantum logic gates^{7–10}.

The physical implementation of quantum networks requires suitable channels and nodes. Photonic channels are well-advanced transmitters of quantum information. Optical photons can carry quantum information over long distances with almost negligible decoherence and are compatible with existing telecommunication fibre technology. The versatility of quantum networks, however, is largely defined by the capability of the network nodes. Dedicated tasks like quantum key distribution can already be achieved using send-only emitter nodes and receive-only detector nodes¹¹. However, in order to fully exploit the capabilities of quantum networks, functional network nodes are required which are able to send, receive and store quantum information reversibly and efficiently.

The implementation and connection of quantum nodes is a major challenge, and different approaches are currently being pursued. An intensely studied example is an ensemble of gas-phase atoms^{12–14}, but the protocols for the generation of single excitations are inherently probabilistic⁹. Another strong contender is a single particle¹⁵, which allows for single-photon emission¹⁶, quantum gate operations^{17–19} and scalability²⁰. But single emitters generally exhibit weak light-matter

interaction resulting, again, in inherently probabilistic information exchange and very low success rates. In particular, the reversible quantum state mapping between a photonic channel and a single emitter in free space is highly inefficient. In their seminal work⁸, Cirac and co-workers therefore proposed to overcome these problems by network nodes based on single emitters embedded in optical cavities.

Here we present the experimental realization of this prototype of a quantum network. The nodes are formed by single atoms quasi-permanently trapped in optical cavities. The cavity-enhanced light-matter interaction opens up a deterministic path for interconversion of photonic and atomic quantum states. By dynamic control of coherent dark states^{8,21–24}, single photons are reversibly exchanged between distant network nodes. We demonstrate faithful quantum state transfer across a network channel and create entanglement between distant nodes. High fidelities and long coherence times are achieved by encoding the quantum information in the polarization of the photon and the atomic ground-state spin. Our results present a direct photonic link between two distant single emitters and pave the way for the realization of large-scale quantum networks.

In the following, we describe experiments in which we characterize a single network node and the connection of two nodes forming an elementary network. The two nodes are operated in independent laboratories at a distance of 21 m and are connected by an optical fibre link of 60 m length. In a first experiment, we demonstrate that single photons can be stored in and retrieved from a single-atom node preserving the photonic polarization state. Second, we show the faithful transfer of arbitrary atomic quantum states from one node to the other. Third, a maximally entangled state of the distant atoms is created with a fidelity of up to 98% and is maintained for at least 100 μ s. This coherence time exceeds the entanglement distribution time across the network link by two orders of magnitude, and translates into a maximum possible entangled node distance of a 20 km optical fibre path. Last, local unitary operations are performed on one of the nodes, resulting in rotations of the non-local bipartite state whereby different maximally entangled states are created.

¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, 85748 Garching, Germany. [†]Present address: Department of Physics, and California NanoSystems Institute, University of California, Santa Barbara, California 93106, USA.

A universal quantum network node

Each node in our quantum network consists of a single neutral rubidium atom that is quasi-permanently trapped in a high-finesse optical resonator^{24,25} (see Methods). The physical parameters put our atom–cavity systems in the intermediate-coupling regime of cavity QED, with the resonators optimized for highly directional optical output into a single mode ($\geq 90\%$) that is efficiently coupled (up to 90%) to the fibre link (Fig. 1). The reversible conversion between quantum states of light and matter is enabled by the dynamic control of coherent dark states. The $5^2S_{1/2}$ hyperfine ground states $F = 1$ and $F = 2$ of the single ^{87}Rb atom are coupled in a Raman configuration formed by the vacuum mode of the cavity (vacuum Rabi frequency $2g$) and an external control laser field (Rabi frequency Ω_C) (Fig. 1). Control laser and cavity are both blue-detuned by several tens of megahertz from the transition to the excited $5^2P_{3/2}$, $F' = 1$ state which—for an ideal coherent dark state—is never populated. The atom is driven coherently between the two ground states by controlling $\Omega_C(t)$, where t is time. With the atom initially prepared in $|F = 2, m_F = 0\rangle$ and $\Omega_C(0) = 0$, a single photon is generated in the cavity mode by increasing the control laser power until $\Omega_C > 2g$. Owing to the finite resonator decay time, the single photon is immediately emitted into the output mode with its temporal wave-packet shape determined by $\Omega_C(t)$ (Fig. 2)²⁶. The single-photon character is confirmed by a measurement of the second-order correlation function $g^{(2)}(\tau)$ (Fig. 2e). The efficiency of the emission process is up to 60%. The inverse process of coherent absorption of a single photon is performed as a time-reversal of the emission process, and requires a decrease in $\Omega_C(t)$ timed to occur when a single photon arrives.

We produce single photons with a nearly time-symmetric envelope in one atom–cavity system (Fig. 2a) and coherently absorb them in the second system⁸. After a selectable storage time, the photon is re-emitted (Fig. 2b). The overall write–read success rate is typically $(10 \pm 1)\%$ (all quoted errors are statistical). With a photon production efficiency of 60%, the storage efficiency is calculated to be 17%. The maximum efficiencies for photon absorption and emission are set by the atom–cavity coupling strength g and can approach unity for smaller cavity mode volumes and vanishing scattering losses at the cavity mirrors.

For encoding one qubit, we utilize the photonic polarization degree of freedom and the atomic ground-state spin, that is, the Zeeman state manifold in each hyperfine state (Fig. 2c and d). When applying a

π -polarized control laser field, the selection rules for electromagnetic dipole transitions ensure the faithful mapping of the polarization state of an incoming photon onto a well-defined superposition of atomic Zeeman states and vice versa for the re-emitted photon. To characterize the conversion process, we set the polarization of the incoming photon and compare it to that of the retrieved photon after storage in the atom²⁴. The reconstructed Poincaré sphere after a storage time of 2.5 μs is shown in Fig. 2f. The average fidelity of the quantum memory, defined as the average overlap with the input state, is $F_{\text{qm}} = (92.2 \pm 0.4)\%$ for photons arriving in a 1 μs time interval (starting at $t = 0.2 \mu\text{s}$ in the graph in Fig. 2b, see Discussion section below). The measured fidelity is far above the classical limit²⁷ of $2/3$. This experiment is the first to prove coherent transfer of a qubit encoded in a single photon onto a single atom. The coherence time of our memory has been characterized earlier and exceeds 100 μs (ref. 24).

Quantum state transfer between single atoms

In the next experiment, we transfer quantum information from node A to the distant node B by sending a single photon across the fibre link. The qubit to be transferred is the state of the atom in node A represented by

$$|\psi_A\rangle = \alpha|F = 1, m_F = -1\rangle + \beta|F = 1, m_F = +1\rangle$$

where α and β are normalized complex-valued amplitudes. We apply a π -polarized control laser pulse to the atom at node A, thereby generating a photon in the polarization state:

$$|\psi_{\text{photon}}\rangle = \alpha|L\rangle + \beta|R\rangle$$

Here $|L\rangle$ and $|R\rangle$ refer to the left and right circular polarization components of the photon. After emission of the photon, the atom at node A is left in the state $|F = 1, m_F = 0\rangle$ (Fig. 3a, left diagram and ref. 21). The photon is transmitted through the optical fibre to node B. The atom at node B is initially prepared in the state $|F = 1, m_F = 0\rangle$. Using the Raman scheme (Fig. 3b, right diagram) the incoming photon is absorbed and its polarization state is mapped onto the atomic state of node B, which becomes:

$$|\psi_B\rangle = \alpha|F = 2, m_F = -1\rangle + \beta|F = 2, m_F = +1\rangle$$

After absorption, an arbitrary quantum state has been successfully communicated from node A to node B. The qubit encoded in atom B

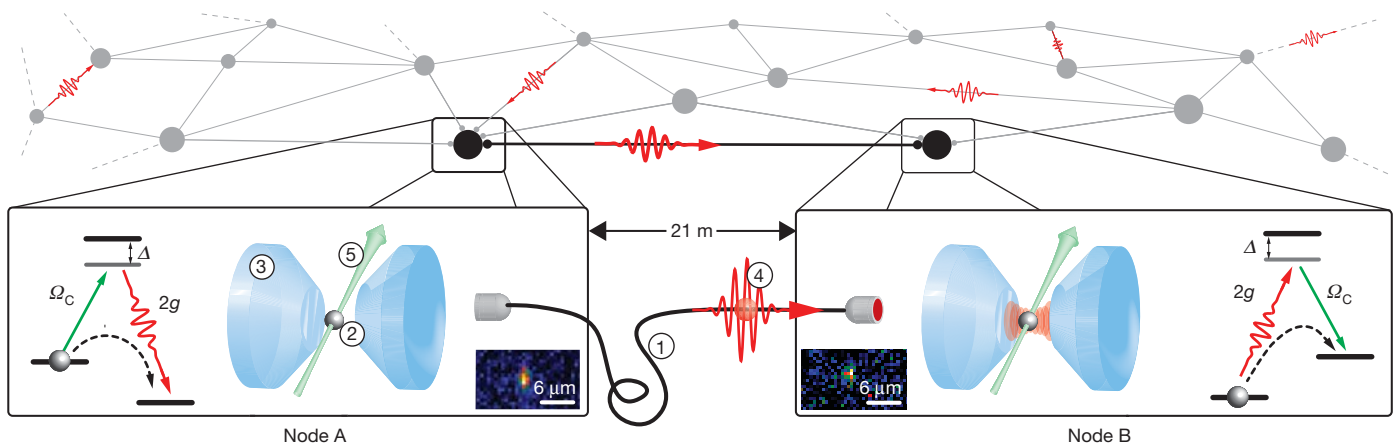


Figure 1 | A cavity-based quantum network. In the envisaged architecture (top), many single-atom nodes are connected by single-photon links. Here we explore the universal properties of the system produced by connecting two nodes (middle; A and B) within this configuration. Details of the nodes and connections are shown in the lower part of the figure. In our experiment, these two identical nodes are located in independent laboratories connected by a 60-m optical fibre (1). Each node consists of a single rubidium atom (2) trapped in an optical dipole trap at the centre of a high-finesse optical cavity (3). Quantum state transfer between the atoms and remote entanglement can be achieved via

exchange of a single photon (4), with the quantum information encoded in the internal state of the atom and the polarization of the photon. Both the production of a photon (node A) and its storage (node B) are achieved via a coherent and reversible stimulated Raman adiabatic passage (see main text for details; (5), control laser). Also shown for each node is the atomic level scheme, with the green and red arrows indicating the control laser and the exchanged single photon, respectively. Δ is the one-photon detuning. Insets, fluorescence images of the two single atoms.

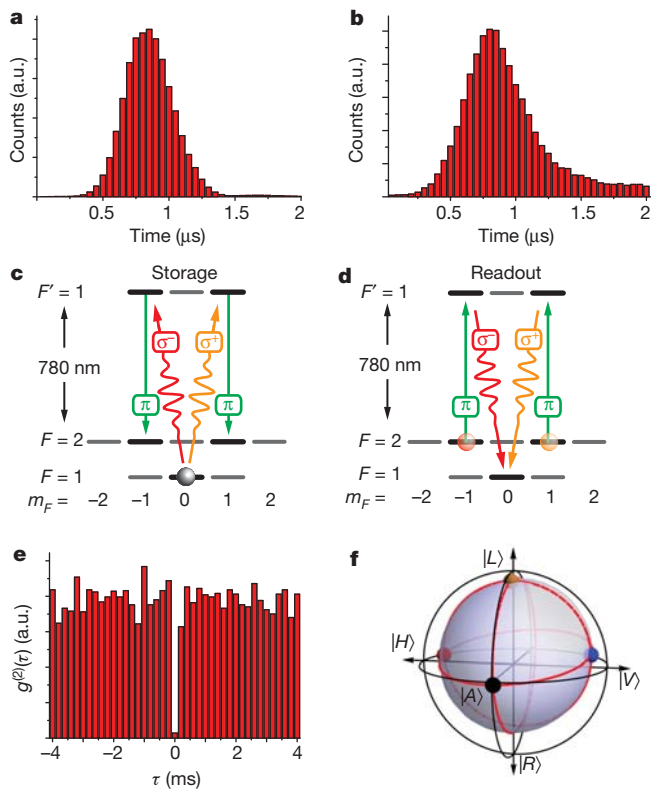


Figure 2 | Universal quantum network node. **a**, Single photons with the indicated temporal envelope (full-width at half-maximum 0.5 μ s) are produced from node A and subsequently stored in node B. **b**, Photons are retrieved after a storage time of 2.5 μ s. **c**, **d**, Atomic level schemes and optical transitions of ^{87}Rb used in the protocol. **c**, The atom at node B is initially prepared in the state $|F=1, m_F=0\rangle$ (grey sphere). The polarization of the incoming photon is a superposition of σ^+ and σ^- components (yellow and red arrows) and is converted into a superposition of the $|F=2, m_F=\pm 1\rangle$ states via a stimulated Raman adiabatic passage using a π -polarized control laser (green arrows). In the process, the phase relation between the σ^\pm polarization components is mapped to a relative phase of the atomic Zeeman states. **d**, The photon is recreated by reversing the storage process, thereby mapping the atomic superposition state (red and yellow sphere) onto the photon's polarization. **e**, Incident photon correlations. The measured second-order correlation function $g^{(2)}(\tau)$ confirms that single photons are produced by node A. **f**, Quantum tomography of the storage process. The minimal deformation of the unit Poincaré sphere proves that every initial photonic quantum state is well preserved during storage and retrieval. The fidelity averaged over all input states is $F_{\text{qm}} = (92.2 \pm 0.4)\%$. $|H\rangle$, $|V\rangle$, $|L\rangle$, $|R\rangle$, $|D\rangle$ and $|A\rangle$ denote horizontal, vertical, left-circular, right-circular, diagonal and anti-diagonal polarization, respectively. All quoted errors are statistical.

is identical to the original one in node A, albeit encoded in the $F=2$ hyperfine manifold. After this quantum state transfer, node A is now ready to receive another photonic qubit, whereas node B is capable of resending the stored qubit at any time. It is this symmetric and reversible feature that makes this scheme scalable to arbitrary network configurations of multiple atom–cavity systems.

We analyse the quantum state transfer using quantum process tomography²⁸. For this purpose, we prepare the atom at node A in a state of the form $|\psi_A\rangle$ by a projective measurement (see Methods). After quantum state transfer from node A to node B as described above, we read out the state of the atom in node B by mapping it onto the polarization of a second single photon which is then detected. By comparing a sufficient set of initial quantum states in node A with the obtained states in node B, we can infer the outcome of the protocol for any initial state of node A. The process matrix χ describes the mapping of the density matrix ρ_A of the state at node A (taken to be the ideally prepared state $|\psi_A\rangle$) onto the transferred state ρ_B at node B

through the operation $\rho_B = \sum_{m,n=0}^3 \chi_{mn} \sigma_m \rho_A \sigma_n^\dagger$. Here, σ_0 is the identity matrix and the other three σ_i are the usual Pauli matrices. For calculating χ we normalize the density matrices.

Ideally, the two density matrices are identical, which is equivalent to having $\chi_{00} = 1$ and all other elements zero. Figure 3c shows the absolute value of all elements of χ obtained from a maximum likelihood fit to the experimental data. We find $\chi_{00} = 0.76$ as the dominating matrix element, indicating a high level of control over the quantum process. The main deviation from a perfect state transfer is a slight depolarization of the quantum state, as indicated by the non-vanishing diagonal elements χ_{11} , χ_{22} and χ_{33} . The state transfer can also be characterized by a fidelity defined as the average overlap between initial and transferred state. We find $F_{\text{trans}} = (84 \pm 1.0)\%$, which proves the quantum character of the state transfer, as the highest fidelity achievable with classical information exchange between the nodes is $2/3$. The overall success rate of the state transfer protocol is 0.2%, resulting from a production efficiency of the transmitter photon at node A of 3% (see Methods), propagation losses leading to a photon transmission of 34% and a storage efficiency at node B of about 20%.

Remote atom–atom entanglement

The most remarkable property of a quantum network is the existence of entangled quantum states shared among several network nodes. This is a basis for quantum logic gate operations between nodes and can lead to complex quantum many-body phenomena.

In the following, we demonstrate the creation of remote entanglement between distant single-atom nodes based on the transmission of a single photon. We first prepare the atom at node A in the state $|F=2, m_F=0\rangle$ (Fig. 4). Applying a π -polarized control laser pulse triggers the emission of a single photon and creates the maximally entangled state

$$|\psi_{A\otimes\text{photon}}^-\rangle = \frac{1}{\sqrt{2}}(|1, -1\rangle \otimes |R\rangle - |1, 1\rangle \otimes |L\rangle)$$

between the spin state of the atom and the polarization of the photon²¹ that is routed to node B. There it is coherently absorbed and its polarization is mapped onto the spin state of atom B. The atom–photon entanglement is thus converted into entanglement between the two nodes, with the two atoms in the maximally entangled $|\Psi^-\rangle$ Bell state:

$$|\psi_{A\otimes B}^-\rangle = \frac{1}{\sqrt{2}}(|1, -1\rangle \otimes |2, 1\rangle - |1, 1\rangle \otimes |2, -1\rangle)$$

We verify the presence of this entangled state by mapping the atomic state at each node onto a photon and analysing the polarization correlations among the two read-out photons. The real part of the resulting density matrix, with the read-out performed 7 μ s after the creation of atom–atom entanglement, is shown in Fig. 4c. We find a fidelity of $F_{|\psi^-\rangle} = (85 \pm 1.3)\%$ with the $|\Psi^-\rangle$ Bell state. This exceeds the classical limit of 50%, clearly proving the existence of entanglement between the two remote atoms. Fidelities as high as $(98.7 \pm 2.2)\%$ can be achieved by further post-selection of photon detection events (see Methods).

The success probability of entanglement creation is 2%. It is the product of the photon generation efficiency (40%) at node A, the probability with which the photon is delivered to node B (34%) and its storage efficiency at node B (14%). The verification process for the entanglement, consisting of the production of one photon at each of the two nodes and their subsequent detection, has an efficiency of 0.16%.

The entanglement created in this experiment exists between the spin states of two single atoms at a physical distance of 21 m. Highly non-classical correlations between the two atoms are observed for 100 μ s. The fidelity with the $|\Psi^-\rangle$ Bell state measured 100 μ s after creation of entanglement is $(56 \pm 3)\%$, still exceeding the classical threshold of 50% by two standard deviations. The decoherence of

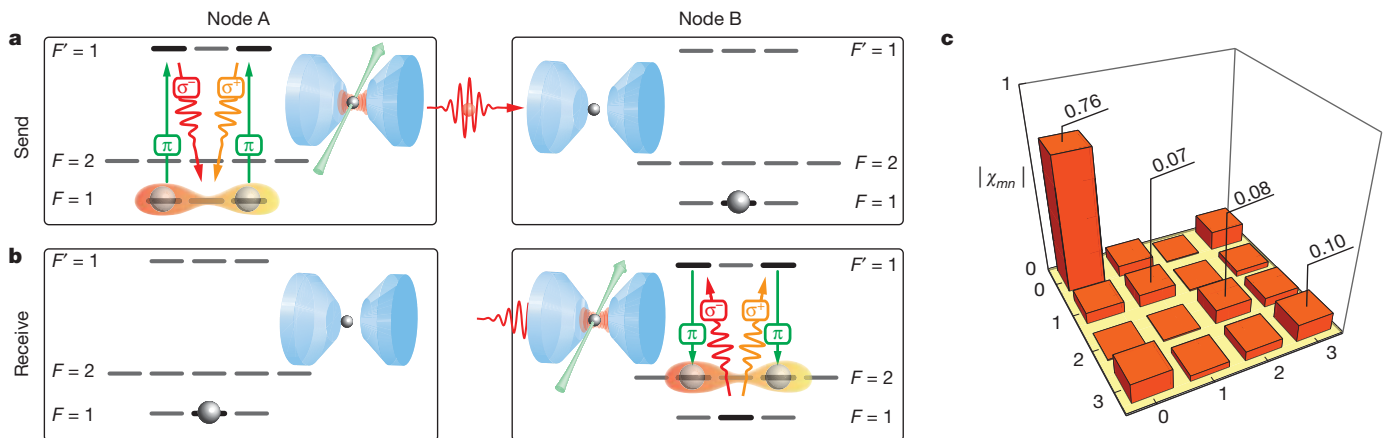


Figure 3 | Quantum state transfer between two single-atom network nodes.

a, At node A (left), an arbitrary quantum state is encoded in the Zeeman state manifold of the single atom (see Methods). This quantum information is mapped onto the polarization of a single photon which is sent to node B (right). **b**, The photonic polarization is mapped to a superposition of atomic Zeeman

states, thereby completing the quantum state transfer from node A to node B. **c**, Absolute value of the elements of the process matrix χ for the quantum state transfer. The average fidelity between the ideal and the read-out transferred state is $(84 \pm 1)\%$, well above the classical limit of $2/3$.

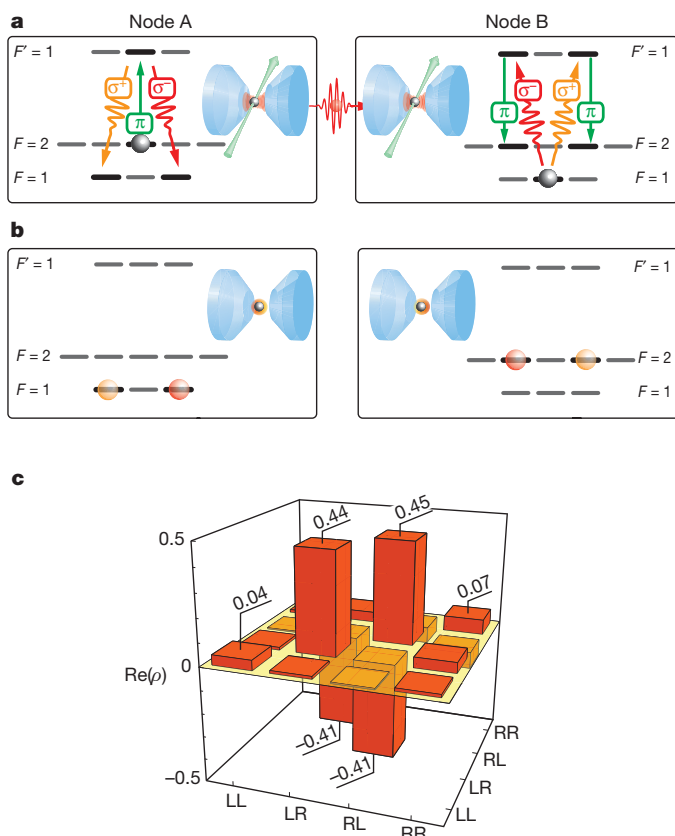


Figure 4 | Remote entanglement of two single-atom nodes. **a**, A single photon is generated at node A (left), such that the internal state of the atom and the polarization of the photon are entangled. The photon is sent to node B (right) where its polarization is mapped onto the atomic state. The grey spheres indicate the initial state of the atoms. **b**, This creates entanglement between nodes A and B that can be maintained for at least $100\ \mu\text{s}$. The atomic levels involved in the entangled state are marked with red and yellow spheres. **c**, For analysis, both atomic states are converted into single photons. Polarization tomography on the two photons confirms the entanglement between the two nodes. We measure a fidelity of $F_{|\Psi^-\rangle} = (85 \pm 1.3)\%$ with respect to the $|\Psi^-\rangle$ Bell state. Shown is the real part of the density matrix. The magnitude of each imaginary part is ≤ 0.03 .

the atom–atom entangled state is dominated by dephasing caused by uncorrelated magnetic field fluctuations (of the order of $1\ \text{mG}$) at the two individual nodes and position-dependent differential a.c. Stark shifts induced by the dipole trap light fields. The dephasing due to local magnetic field fluctuations can be significantly reduced by applying small magnetic guiding fields ($30\ \text{mG}$) along the quantization axis of each node. This has been used for the measurements with $100\ \mu\text{s}$ entanglement duration. The observed remote-entanglement lifetime exceeds the entanglement creation time ($1\ \mu\text{s}$ for creation, transmission and absorption of an entangling photon) by two orders of magnitude. Entanglement lifetimes of the order of seconds can be expected when mapping the Zeeman qubit to magnetic-field-insensitive clock states using microwave or Raman pulses²⁹.

In the limit of unit efficiency, the entanglement scheme presented here allows for the deterministic creation of entanglement. In our experimental implementation, efficiencies are below one and we therefore detect a posteriori entanglement³⁰. The detection of entangled read-out photons indicates that atom–atom entanglement had been present. Only entanglement attempts that lead to the final detection of two read-out photons in the mapping process are considered in our data. The creation of heralded entanglement^{31–33} is possible by implementing a mechanism that signals the successful storage of a transmitted photon at node B without disturbing the stored quantum state (see Discussion below).

Local manipulation of a non-local state

Nodes A and B are in separated physical locations and thus are independently addressable for local qubit control. When two nodes are entangled, unitary operations applied locally at one of the nodes change the non-local state of both nodes while the entanglement is preserved. Thus local qubit control allows arbitrary maximally entangled two-qubit states to be created using a single initial entangled state as a resource. We demonstrate this capability by creating the $|\Psi^+\rangle$ Bell state. We start by preparing the two nodes in the $|\Psi^-\rangle$ Bell state as described above. Applying a magnetic field along the quantization axis only at node B causes a state rotation at twice the Larmor frequency. The fidelity of the created state with the $|\Psi^-\rangle$ and the $|\Psi^+\rangle$ Bell state is plotted as a function of the applied magnetic field in Fig. 5. The time between entanglement creation and read-out of the atomic state is fixed at $12.5\ \mu\text{s}$. As can be seen from Fig. 5, the rotation of the non-local state results in a sinusoidally varying overlap with the $|\Psi^\pm\rangle$ Bell states. The fidelity with respect to the $|\Psi^+\rangle$ state reaches a maximum of $(81 \pm 2)\%$ at a magnetic field of $B = 30\ \text{mG}$. The original $|\Psi^-\rangle$ state is recovered with a fidelity of $(76 \pm 2)\%$ after a spin

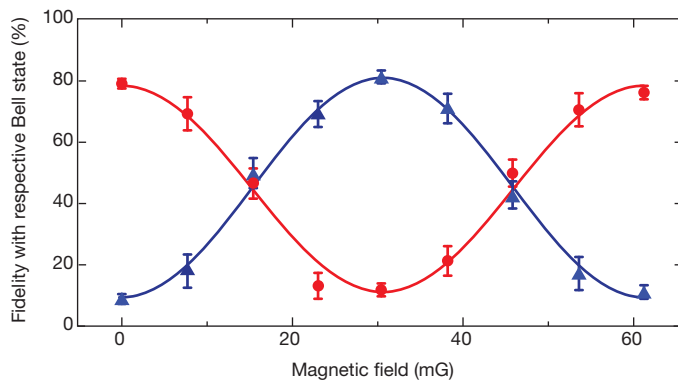


Figure 5 | Controlled rotation of the entangled state. A magnetic field, locally applied at node B, changes the non-local quantum state of the entangled nodes. The atom at node A is held at zero magnetic field. The phase evolution for a fixed hold time of 12.5 μ s is proportional to the applied magnetic field. Plotted are the fidelity with the $|\Psi^- \rangle$ and the $|\Psi^+ \rangle$ Bell state (red and blue data, respectively), showing a characteristic oscillation. The error bars indicate the statistical standard error. The solid line is a cosine fit to guide the eye. The initially prepared $|\Psi^- \rangle$ state ($F_{|\Psi^- \rangle} = (79 \pm 1.6)\%$) can be rotated into a $|\Psi^+ \rangle$ state of comparable fidelity ($F_{|\Psi^+ \rangle} = (81 \pm 2)\%$) using a magnetic field of 30 mG.

rotation of 2π . The reduced fidelity with the $|\Psi^- \rangle$ state at a field of 60 mG is a result of non-negligible Larmor precession during the entanglement creation and read-out processes.

Discussion

The body of work presented here constitutes, to our knowledge, the first direct coupling of two distant single quantum emitters by exchange of a single photon. Our results introduce universal quantum network nodes based on single emitters. Single-atom-cavity nodes excel previously investigated light-matter interfaces and incorporate their specific advantages in one platform. The use of single emitters offers a clear perspective for heralding^{34–36} and the integration of quantum gate operations both local and remote^{15,17–19,37}. In the following, we briefly discuss the potential of our specific implementation of universal quantum network nodes with respect to fidelity and efficiency of the described processes, storage time and scalability.

In all our experiments, atomic state preparation errors due to non-optimal optical pumping reduce the fidelity. Control lasers that are not perfectly π -polarized and off-resonant excitations cause deviations from the ideal transition scheme (Figs 2–4). These errors may lead to emission of photons with excitation paths different from the Raman scheme. A detailed analysis of transition strengths and effective Rabi frequencies shows that these excitation paths lead to delayed photon emission. The contribution of these photons to the measured signal can therefore be suppressed by post-selecting subsets of data based on photon arrival times (see Methods). As an example, entangled-state fidelities as high as $(98.7 \pm 2.2)\%$ are reached when considering only the first 50% of read-out photons from node A and the initial 14% of photons retrieved from node B. These results show the great potential for achieving very high fidelities if the mentioned imperfections are overcome.

The efficiency achievable with the demonstrated deterministic entanglement scheme⁸ is higher than what can be achieved with probabilistic schemes^{9,15}. Although our first implementation is not deterministic and the entanglement is not heralded, our efficiencies exceed previous demonstrations of remote entanglement by several orders of magnitude^{31–33}. The main limitation in our experiments is the moderate atom-cavity coupling strength, g . Efficiencies approaching unity may be accomplished when the cavity mode volume is decreased³⁶ and mirror scattering losses are eliminated. Ultimately, the loss of photons in the optical fibre connecting two quantum nodes will limit the overall efficiency for long distances, underlining the importance of a heralding scheme.

Heralding is crucial in networks with non-unity success probability of the quantum link, and ensures scalability in more complex and long-distance quantum networks. For architectures based on single atom-cavity nodes this is conceptually straightforward. The successful storage process is associated with a change of the atomic hyperfine ground state. The depopulation of the initial state can then be probed spectroscopically, using, for example, cavity-assisted fluorescence hyperfine state detection or cavity transmission^{34–36}. As the splitting between the two hyperfine ground states in ^{87}Rb is large, optical probing of the $F=1$ population will leave the qubit stored in $F=2$ unaltered. The main challenge is to detect $F=1$ with high fidelity before off-resonant excitation leads to decay into $F=2$, that is, with only a few scattered photons. Although the implementation of heralding is therefore challenging with our cavities, it is certainly feasible with current cavity technology³⁶.

The very weak coupling of the nuclear spin of single atoms to the environment can be exploited to boost the coherence times of our network nodes by mapping the Zeeman qubit onto magnetic-field insensitive clock states²⁹. The number of qubits per node may be increased through the use of optical lattices and single-atom registers³⁸. After the preparation of remote atom-atom entanglement using a procedure akin to that shown here, the registers in different nodes could be shifted to successively produce many sets of entangled atoms which can then be used, for example, for nested entanglement purification. This possibility, in combination with the long storage times achievable with single atoms and the potential for heralding, represents a realistic avenue towards quantum communication over arbitrary distances by means of a quantum repeater protocol¹⁰.

Current advances in photonic technologies allow reconfigurable routing between different nodes, thereby enabling various different network topologies. The controlled interaction between arbitrary nodes and the plethora of accessible topologies of cavity quantum networks are not only an important resource for quantum information processing; cavity networks also constitute a suitable paradigm for investigating emergent phenomena, such as quantum phase transitions of light^{4–6} or percolation of entanglement¹.

METHODS SUMMARY

The two independent quantum nodes are designed to operate with similar physical parameters. In each apparatus, a single ^{87}Rb atom is quasi-permanently trapped inside an optical dipole trap (potential depth $U_0/k_B = 3$ and 5 mK, where k_B is the Boltzmann constant) and held at the centre of a high-finesse optical cavity (finesse 6×10^4 , mirror distance 0.5 mm, mode waist 30 μ m). Both cavities have asymmetric mirror transmissions of $T_1 < 6$ p.p.m. and $T_2 \approx 100$ p.p.m., leading to a highly directional ($T_{\text{out}} \geq 0.9$) single output mode. Both systems produce photons on the D_2 line of ^{87}Rb at a wavelength of 780 nm. In this configuration, both atom-cavity systems operate in the intermediate-coupling regime of cavity QED (coherent atom-cavity coupling $g \leq 2\pi \times 5$ MHz, cavity field decay rate $\kappa = 2\pi \times 3$ MHz, atomic polarization decay rate $\gamma = 2\pi \times 3$ MHz). When a single atom is trapped inside each of the cavities, the experimental protocol runs at a repetition rate of 5 kHz, including optical pumping (20 μ s), photon generation (1 μ s), photon storage (1–100 μ s) and optical cooling of atomic motion (80 μ s). The necessary laser beams impinge perpendicular to the cavity axis. The presence and position of single atoms is monitored in real time by an electron multiplying CCD camera, which collects atomic fluorescence light (Fig. 1 inset). In combination with a longitudinally shiftable standing-wave dipole trap, the atoms are actively positioned at the centre of the cavity mode³⁸. Single-atom storage times are of the order of one minute.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 December 2011; accepted 6 March 2012.

1. Acín, A., Cirac, J. I. & Lewenstein, M. Entanglement percolation in quantum networks. *Nature Phys.* **3**, 256–259 (2007).
2. Choi, K. S., Goban, A., Papp, S. B., van Enk, S. J. & Kimble, H. J. Entanglement of spin waves among four quantum memories. *Nature* **468**, 412–416 (2010).
3. Jungnitsch, B., Moroder, T. & Gühne, O. Taming multiparticle entanglement. *Phys. Rev. Lett.* **106**, 190502 (2011).
4. Törmä, P. Transitions in quantum networks. *Phys. Rev. Lett.* **81**, 2185–2189 (1998).

5. Hartmann, M. J., Brandão, F. G. S. L. & Plenio, M. B. Strongly interacting polaritons in coupled arrays of cavities. *Nature Phys.* **2**, 849–855 (2006).
6. Greentree, A. D., Tahan, C., Cole, J. H. & Hollenberg, L. C. L. Quantum phase transitions of light. *Nature Phys.* **2**, 856–861 (2006).
7. Kimble, H. J. The quantum internet. *Nature* **453**, 1023–1030 (2008).
8. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.* **78**, 3221–3224 (1997).
9. Duan, L.-M., Lukin, M. D., Cirac, J. I. & Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414**, 413–418 (2001).
10. Briegel, H.-J., Dür, W., Cirac, J. I. & Zoller, P. Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998).
11. Eisaman, M. D., Fan, J., Migdall, A. & Polyakov, S. V. Invited review article: Single-photon sources and detectors. *Rev. Sci. Instrum.* **82**, 071101 (2011).
12. Lvovsky, A. I., Sanders, B. C. & Tittel, W. Optical quantum memory. *Nature Photon.* **3**, 706–714 (2009).
13. Hammerer, K., Sørensen, A. S. & Polzik, E. S. Quantum interface between light and atomic ensembles. *Rev. Mod. Phys.* **82**, 1041–1093 (2010).
14. Sangouard, N., Simon, C., de Riedmatten, H. & Gisin, N. Quantum repeaters based on atomic ensembles and linear optics. *Rev. Mod. Phys.* **83**, 33–80 (2011).
15. Duan, L.-M. & Monroe, C. Colloquium: Quantum networks with trapped ions. *Rev. Mod. Phys.* **82**, 1209–1224 (2010).
16. Lounis, B. & Orrit, M. Single-photon sources. *Rep. Prog. Phys.* **68**, 1129–1179 (2005).
17. Jaksch, D. *et al.* Fast quantum gates for neutral atoms. *Phys. Rev. Lett.* **85**, 2208–2211 (2000).
18. Isenhower, L. *et al.* Demonstration of a neutral atom controlled-NOT quantum gate. *Phys. Rev. Lett.* **104**, 010503 (2010).
19. Timoney, N. *et al.* Quantum gates and memory using microwave-dressed states. *Nature* **476**, 185–188 (2011).
20. Home, J. P. *et al.* Complete methods set for scalable ion trap quantum information processing. *Science* **325**, 1227–1230 (2009).
21. Wilk, T., Webster, S. C., Kuhn, A. & Rempe, G. Single-atom single-photon quantum interface. *Science* **317**, 488–490 (2007).
22. Boozer, A. D., Boca, A., Miller, R., Northup, T. E. & Kimble, H. J. Reversible state transfer between light and a single trapped atom. *Phys. Rev. Lett.* **98**, 193601 (2007).
23. Mücke, M. *et al.* Electromagnetically induced transparency with single atoms in a cavity. *Nature* **465**, 755–758 (2010).
24. Specht, H. P. *et al.* A single-atom quantum memory. *Nature* **473**, 190–193 (2011).
25. Lettner, M. *et al.* Remote entanglement between a single atom and a Bose-Einstein condensate. *Phys. Rev. Lett.* **106**, 210503 (2011).
26. Kuhn, A., Hennrich, M. & Rempe, G. Deterministic single-photon source for distributed quantum networking. *Phys. Rev. Lett.* **89**, 067901 (2002).
27. Massar, S. & Popescu, S. Optimal extraction of information from finite quantum ensembles. *Phys. Rev. Lett.* **74**, 1259–1263 (1995).
28. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
29. Olmschenk, S. *et al.* Manipulation and detection of a trapped Yb⁺ hyperfine qubit. *Phys. Rev. A* **76**, 052314 (2007).
30. van Enk, S. J., Lütkenhaus, N. & Kimble, H. J. Experimental procedures for entanglement verification. *Phys. Rev. A* **75**, 052318 (2007).
31. Moehring, D. L. *et al.* Entanglement of single-atom quantum bits at a distance. *Nature* **449**, 68–71 (2007).
32. Chou, C.-W. *et al.* Functional quantum nodes for entanglement distribution over scalable quantum networks. *Science* **316**, 1316–1320 (2007).
33. Yuan, Z.-S. *et al.* Experimental demonstration of a BDCZ quantum repeater node. *Nature* **454**, 1098–1101 (2008).
34. Lloyd, S., Shahriar, M. S., Shapiro, J. H. & Hemmer, P. R. Long distance, unconditional teleportation of atomic states via complete Bell state measurements. *Phys. Rev. Lett.* **87**, 167903 (2001).
35. Bochmann, J. *et al.* Lossless state detection of single neutral atoms. *Phys. Rev. Lett.* **104**, 203601 (2010).
36. Volz, J., Gehr, R., Dubois, G., Estève, J. & Reichel, J. Measurement of the internal state of a single atom without energy exchange. *Nature* **475**, 210–213 (2011).
37. Duan, L.-M., Wang, B. & Kimble, H. J. Robust quantum gates on neutral atoms with cavity-assisted photon scattering. *Phys. Rev. A* **72**, 032333 (2005).
38. Nußmann, S. *et al.* Submicron positioning of single atoms in a microcavity. *Phys. Rev. Lett.* **95**, 173602 (2005).

Acknowledgements We thank D. Moehring for contributions during the early stage of the experiments, and B. Mayer and M. Padilla from the Walter Schottky Institut for gold coating of the fast-moving mirror. This work was supported by the Deutsche Forschungsgemeinschaft (Research Unit 635), by the European Union (Collaborative Project AQUITE) and by the Bundesministerium für Bildung und Forschung via IKT 2020 (QK_QuOREP). E.F. acknowledges support from the Alexander von Humboldt Foundation.

Author Contributions All authors contributed to the experiment, the analysis of the results and the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.R. (stephan.ritter@mpq.mpg.de).

METHODS

Experimental set-up. The two independent quantum nodes are designed to operate with similar physical parameters. In each apparatus, a single ^{87}Rb atom is quasi-permanently trapped inside an optical dipole trap (potential depth $U_0/k_B = 3$ and 5 mK, where k_B is the Boltzmann constant) and held at the centre of a high-finesse optical cavity (finesse 6×10^4 , mirror distance 0.5 mm, mode waist 30 μm). Both cavities have asymmetric mirror transmissions of $T_1 < 6$ p.p.m. and $T_2 \approx 100$ p.p.m., leading to a highly directional ($T_{\text{out}} \geq 0.9$) single output mode. Both systems produce photons on the D_2 line of ^{87}Rb at a wavelength of 780 nm. In this configuration both atom–cavity systems operate in the intermediate-coupling regime of cavity QED (coherent atom–cavity coupling $g \leq 2\pi \times 5$ MHz, cavity field decay rate $\kappa = 2\pi \times 3$ MHz, atomic polarization decay rate $\gamma = 2\pi \times 3$ MHz). When a single atom is trapped inside each of the cavities, the experimental protocol runs at a repetition rate of 5 kHz, including optical pumping (20 μs), photon generation (1 μs), photon storage (1–100 μs) and optical cooling of atomic motion (80 μs). The necessary laser beams impinge perpendicular to the cavity axis. The presence and position of single atoms is monitored in real time by an electron multiplying CCD camera, which collects atomic fluorescence light (Fig. 1 inset). In combination with a longitudinally shiftable standing-wave dipole trap, the atoms are actively positioned at the centre of the cavity mode³⁸. Single-atom storage times are of the order of one minute.

Projective atomic state preparation for quantum state transfer. To characterize the quality of the quantum state transfer from node A to node B, the atom at node A is prepared in one of six different initial states of the form $|\psi_A\rangle = \alpha|F=1, m_F=-1\rangle + \beta|F=1, m_F=+1\rangle$, forming a regular octahedron on the Poincaré sphere. This is achieved through a projective measurement. The atom is initialized in the state $|F=2, m_F=0\rangle$, and subsequently a single photon is generated using a π -polarized laser pulse. This creates the atom–photon entangled state $|\psi_{A\otimes\text{photon}}\rangle = \frac{1}{\sqrt{2}}(|1,-1\rangle \otimes |R\rangle - |1,1\rangle \otimes |L\rangle)$. Detection of this photon in a well-defined polarization state projects the atom in node A onto a qubit state $|\psi_A\rangle$, with (α, β) determined by the particular choice of the detector's polarization basis. Following this projective preparation, the known quantum state of node A is transferred to node B. Read-out of the state of the atom at node B is performed by mapping it onto a single photon whose polarization can be analysed. In calculating the fidelity of the quantum process, we assume perfect preparation of $|\psi_A\rangle$. It is therefore a lower bound on the fidelity of the state transfer.

Post-selected fidelities. The ideal Raman schemes depicted in Figs 2–4 lead to emission of spatiotemporally well-defined single-photon wave packets with their polarization determined by the selection rules. The fidelities reported in the previous sections are obtained from analysing correlations between photon detection events in different polarization bases. We have identified several experimental imperfections which cause deviations from this ideal Raman scheme: imperfect initial state preparation, misaligned polarization of the control laser and off-resonant excitations. These imperfections not only affect polarization correlations but also the temporal wave packet shape of the emitted photons. A detailed analysis has shown that these imperfections are generally correlated with delayed photon emission with respect to the ideal Raman scheme. In this Article, we usually evaluate all read-out photons from node A and those photons from node B arriving within a 1 μs time interval centred around the maximum of the photon wave packet (see Fig. 2b). The contribution of the mentioned non-ideal processes to the measured fidelities can however be minimized when only those photon detection events are analysed that occur early in the photon's temporal wave packet. Indeed, we find close to ideal fidelities for these subsets of data. In the atom–atom entanglement experiment the fidelity is increased to

$F_{|\psi\rangle} = (98.7 \pm 2.2)\%$ when considering only the first 50% of the ensemble of detected photons from node A and the initial 14% of the ensemble of photons from node B. The quoted values for $F_{|\psi\rangle}$ are the unbiased estimator and the statistical standard error. The likelihood function of $F_{|\psi\rangle}$ is non-Gaussian.

The imperfections correlated with late photon emission can also be suppressed by tailoring the single-photon read-out process directly. When the applied control laser Rabi frequencies Ω_C are kept low and applied for only a short time, the read-out photon can be made to resemble the post-selected subset. We have made use of this weak read-out in the experiments on state transfer and remote atom–atom entanglement, thereby optimizing for high fidelities at the expense of the efficiency of the process. As all presented schemes are intrinsically deterministic, future improvements of the imperfect processes mentioned above will allow for both fidelities and efficiencies approaching unity, without the necessity for any trade-off.

Success probabilities. Single photons are produced with an efficiency $\eta_{F=2}^B = 0.6$ at node B, and $\eta_{F=2}^A = 0.4$ and $\eta_{F=1}^A = 0.03$ at node A. The indices $F=1$ and $F=2$ denote the initial hyperfine state of the atom. $\eta_{F=1}^A$ is deliberately kept low to suppress off-resonant excitations to nearby hyperfine states. As this is only relevant for atoms prepared in $F=1$ because of the near-resonant cavity, it can be circumvented by a local transfer of the qubit from the $F=1$ to the $F=2$ manifold using optical Raman or microwave pulses. The storage efficiency was $S_{\text{trans}} = 0.2$ and $S_{\text{ent}} = 0.14$ in the experiments on quantum state transfer and remote atom–atom entanglement, respectively. An intracavity photon is outcoupled into one well-defined free-space mode with a probability of $T_{\text{out}} = 0.9$, which is matched to a single-mode optical fibre (fibre coupling efficiency up to 0.9). A fast-moving mirror switches the network between two configurations. The network's nodes are either directly connected through an optical fibre path with a transmission $T_{\text{net}} = 0.4$ (measured from before fibre input at node A to fibre output at node B). Or the nodes are disconnected and photons emanating from each node are guided to separate polarization detection set-ups with an optical path transmission of $T_{\text{det}} = 0.6$. With our photodetector quantum efficiency $\varepsilon = 0.6$, this yields a detection probability of $P_{\text{det}} = T_{\text{out}}T_{\text{det}}\varepsilon = 0.3$ for a given intracavity photon.

From these numbers the success probability for quantum state transfer can be calculated: $P_{\text{trans}} = \eta_{F=1}^A T_{\text{out}} T_{\text{net}} S_{\text{trans}} = 0.2\%$. The efficiency of the projective atomic state preparation at node A is $P_{\text{prep}} = \eta_{F=2}^A P_{\text{det}} = 12\%$ and the experiments are repeated at a rate of $f = 5$ kHz. The duty cycle D denotes the fraction of the total measurement time during which data were taken. One limiting factor is the condition that a single atom needs to be trapped and localized at the centre of the cavity mode of node A and B simultaneously. The duty cycle is $D_{\text{trans}} = 0.2$ and $D_{\text{ent}} = 0.3$ in the experiments on quantum state transfer and remote atom–atom entanglement, respectively. This yields a rate for the state transfer of $R_{\text{trans}} = P_{\text{trans}}/D_{\text{trans}} \approx 2$ per second. The rate of successfully verified attempts, each consisting of quantum state preparation at node A, quantum state transfer and detection of the readout photon from node B, is $R_{\text{trans}}^{\text{det}} = P_{\text{prep}} P_{\text{trans}} \eta_{F=2}^B P_{\text{det}}/D_{\text{trans}} \approx 3$ per minute.

The success probability for remote entanglement creation is $P_{\text{ent}} = \eta_{F=2}^A T_{\text{out}} T_{\text{net}} S_{\text{ent}} = 2\%$, yielding a rate of $R_{\text{ent}} = P_{\text{ent}}/D_{\text{ent}} \approx 30$ entanglement creations per second and $R_{\text{ent}}^{\text{det}} = P_{\text{ent}} \eta_{F=1}^A P_{\text{det}} \eta_{F=2}^B P_{\text{det}}/D_{\text{ent}} \approx 3$ detected events of a posteriori entanglement per minute. Direct readout of the atomic states with near-unity efficiency³¹ would increase the detected entanglement event rate to the value of R_{ent} . Beyond the potential for higher efficiencies via an increased atom–cavity coupling strength (see Discussion), a higher rate can be achieved in steeper trapping potentials that allow for more efficient and therefore faster cooling.

Teneurins instruct synaptic partner matching in an olfactory map

Weizhe Hong¹, Timothy J. Mosca¹ & Liqun Luo¹

Neurons are interconnected with extraordinary precision to assemble a functional nervous system. Compared to axon guidance, far less is understood about how individual pre- and postsynaptic partners are matched. To ensure the proper relay of olfactory information in the fruitfly *Drosophila*, axons of ~50 classes of olfactory receptor neurons (ORNs) form one-to-one connections with dendrites of ~50 classes of projection neurons (PNs). Here, using genetic screens, we identified two evolutionarily conserved, epidermal growth factor (EGF)-repeat containing transmembrane Teneurin proteins, Ten-m and Ten-a, as synaptic-partner-matching molecules between PN dendrites and ORN axons. Ten-m and Ten-a are highly expressed in select PN-ORN matching pairs. Teneurin loss- and gain-of-function cause specific mismatching of select ORNs and PNs. Finally, Teneurins promote homophilic interactions *in vitro*, and Ten-m co-expression in non-partner PNs and ORNs promotes their ectopic connections *in vivo*. We propose that Teneurins instruct matching specificity between synaptic partners through homophilic attraction.

The chemoaffinity hypothesis was proposed nearly 50 years ago to explain the target specificity of regenerating optic nerves: developing neurons “must carry individual identification tags, presumably cytochemical in nature, by which they are distinguished one from another almost, in many regions, to the level of the single neuron”¹. Many molecules are now known that guide axons to their target areas^{2,3}, but few may mediate mutual selection and direct matching between individual pre- and postsynaptic partners. Here we show that the transmembrane Teneurin proteins instruct the selection of specific synaptic partners in the *Drosophila* olfactory circuit (Supplementary Fig. 1).

In *Drosophila*, individual classes of ORN axons make one-to-one connections with individual classes of second-order PN dendrites within one of ~50 discrete glomeruli in the antennal lobe. We refer to this specific one-to-one connection as PN-ORN synaptic partner matching. Olfactory circuit assembly takes place in sequential steps before sensory activity begins^{4–6}. PN dendrites first elaborate within and pattern the developing antennal lobe^{7–9}, which is followed by invasion of ORN axons^{10–14}. Importantly, re-positioning PN dendrites redirects their partner ORN axons without disrupting the connections¹⁵, suggesting that proper PN-ORN connections probably involve direct recognition and matching between partners.

Matching screens identified Ten-m and Ten-a

To identify potential PN-ORN matching molecules, we simultaneously labelled select PN dendrites and ORN axons in two colours and performed two complementary genetic screens (Fig. 1a, d). We overexpressed 410 candidate cell-surface molecules, comprising ~40% of the potential cell-recognition molecules in *Drosophila*¹⁶. In the first screen, we used *Mz19-GAL4* to label DA1, VA1d and DC3 PNs (hereafter Mz19 PNs), and *Or47b-rCD2* to label Or47b ORNs (Fig. 1a, b). Or47b ORN axons normally project to the VA1m glomerulus and are adjacent to Mz19 PN dendrites without overlap. We overexpressed candidate cell-surface molecules only in Mz19 PNs to identify those that promoted ectopic connections between Or47b axons and Mz19 dendrites (Fig. 1a). We found that overexpression of *ten-m* (*P{GS}9267*; Supplementary Fig. 2b) produced ectopic connections (Fig. 1c).

In the second screen, we labelled Mz19 PNs as above and Or88a ORNs using *Or88a-rCD2* (Fig. 1d, e). Or88a ORN axons normally project to the VA1d glomerulus, intermingling extensively with VA1d PN dendrites (Fig. 1e). We overexpressed candidate cell-surface molecules in Mz19 PNs (Fig. 1d) as above and found that overexpression of *ten-a* (*P{GE}1914*, Supplementary Fig. 2a) partially disrupted the intermingling of Or88a axons and Mz19 dendrites (Fig. 1f).

In addition to impairing PN-ORN matching, *ten-m* and *ten-a* overexpression shifted Mz19 PN dendrite position (Fig. 1c, f). However, mismatching was not a secondary consequence of axon or dendrite mispositioning; mispositioning alone, caused by perturbation of other genes, does not alter PN-ORN matching^{9,13,15}. Furthermore, among 410 candidate molecules, only *ten-m* and *ten-a* overexpression exhibited mismatching defects, suggesting their specificity in PN-ORN matching.

Both *ten-m* and *ten-a* appear to encode type II transmembrane proteins^{17–19}. They possess highly similar domain compositions and amino acid sequences; each contains eight EGF-like and multiple YD (tyrosine-aspartate) repeats within its large carboxy-terminal extracellular domain (Fig. 1g). Ten-m and Ten-a were initially identified as tenascin-like molecules^{20,21}, but vertebrate teneurins were later identified as their true homologues based on sequence and domain similarity (Fig. 1h). Thus, we refer to Ten-m and Ten-a as *Drosophila* Teneurins. Teneurins are present in nematodes, flies and vertebrates. In human, teneurin-1 and teneurin-2 are located in chromosomal regions associated with intellectual disability¹⁷, and teneurin-4 is linked to susceptibility to bipolar disorder²².

Drosophila ten-m was originally identified as a pair-rule gene required for embryonic patterning^{21,23}, but this function was recently shown to be unrelated to *ten-m*²⁴. Teneurins were implicated in synapse development at the neuromuscular junction^{16,25} (see ref. 26), and Ten-m also regulates motor axon guidance²⁴. Neither the underlying mechanisms nor their potential roles in the central nervous system are known. Vertebrate teneurins are widely expressed in the nervous system^{18,27} and interact homophilically *in vitro*^{28,29}, suggesting their potential role as homophilic cell adhesion molecules in patterning neuronal connectivity.

¹Department of Biology, Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA.

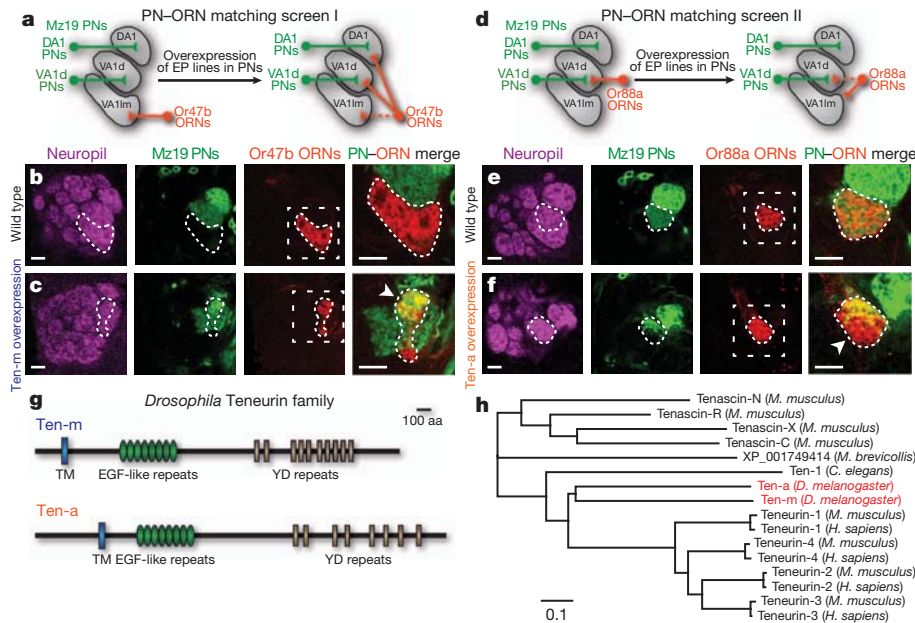


Figure 1 | PN-ORN synaptic matching screens identify two Teneurins.

a, d, Schematics showing two PN-ORN matching screens. PN dendrites are labelled by *Mz19-GAL4* driving mCD8GFP and ORN axons by *Or47b-rCD2* (a) or *Or88a-rCD2* (d). Each EP line has a transposable element insertion that places *UAS* 5' to a gene encoding a predicted cell-surface protein, which can be overexpressed using *Mz19-GAL4*. **b, c,** Or47b axons and Mz19 dendrites do not overlap in control (b), but form ectopic connections following *Ten-m* overexpression (c), as seen by axon-dendrite intermingling (arrowhead). **e, f,** Or88a axons and Mz19 dendrites connect at the VA1d glomerulus in control (e), but the connection is partially lost following *Ten-a* overexpression, as some of the Or88a axons no longer intermingle with Mz19 dendrites

Matching expression of *Ten-m* and *Ten-a*

Both *Drosophila* Teneurin proteins were endogenously expressed in the developing antennal lobe (Fig. 2a and Supplementary Fig. 3). At 48 h after puparium formation (APF), when individual glomeruli just become identifiable, elevated Teneurin expression was evident in select glomeruli. The subset of glomeruli expressing elevated *Ten-m* was distinct but partially overlapping with that expressing elevated *Ten-a* (Fig. 2a, e). Teneurin proteins were also detected at a low level in all glomeruli. Both basal and elevated Teneurin expressions were eliminated by pan-neuronal RNA interference (RNAi) targeting the corresponding gene (Fig. 2b, c), suggesting that Teneurin proteins are produced predominantly by neurons. In a *ten-a* null mutant we generated (Supplementary Fig. 2a), all *Ten-a* expression was eliminated, confirming antibody specificity (Fig. 2d).

The antennal lobe consists of ORN axons as well as PN and local interneuron dendrites. We used intersectional analysis to determine the cellular source for elevated Teneurin expression. For *ten-m*, we screened *GAL4* enhancer traps near the *ten-m* gene, and identified NP6658 (hereafter *ten-m-GAL4*; Supplementary Fig. 2b) that recapitulated the glomerulus-specific *Ten-m* staining pattern (Supplementary Fig. 4a–c). We used a FLPout reporter *UAS>stop>mCD8GFP* to determine the intersection of *ten-m-GAL4* and an ORN-specific *ey-Flp* (Fig. 2f and Supplementary Fig. 4d–f) or a PN-specific *GH146-Flp* (Fig. 2g and Supplementary Fig. 4g–i). We found that *ten-m-GAL4* was selectively expressed in a subset of ORNs and PNs. Owing to reagent availability, we focused our analysis on five glomeruli (DA1, VA1d, VA1lm, DC3 and DA3), adjacently located on the lateral and anterior side of the antennal lobe. In these five glomeruli, *Ten-m* expression in PN and ORN classes matched: high levels in PNs corresponded to high levels in ORNs and vice versa (Fig. 2f, g).

To determine the cellular origin of elevated *Ten-a* expression, we performed tissue-specific RNAi of endogenous *Ten-a*, as no *GAL4*

(arrowhead). Target areas of Or47b (b, c) or Or88a (e, f) axons are outlined. Mismatching phenotypes are quantified in Supplementary Figs 9k and 10q. The first three columns in b, c, e, f show separate channels of the same section; the fourth shows higher magnification of the dashed squares (as in Figs 3, 4, 5d–g). Unless indicated, all images in this and subsequent figures are single confocal sections and all scale bars are 10 μ m. **g,** Domain composition of *Drosophila* *Ten-m* and *Ten-a*. aa, amino acids; TM, transmembrane domain. **h,** Phylogeny of the *Drosophila* Teneurins and related proteins in other species. Branch lengths represent units of substitutions per site of the sequence alignment. Teneurins are evolutionarily conserved in bilaterians and a unicellular choanoflagellate *Monosiga brevicollis*, but not in cnidarians.

enhancer trap is available near *ten-a*. To isolate *Ten-a* expression in ORNs, we drove pan-neuronal *ten-a* RNAi while specifically suppressing RNAi in ORNs using *tubP>stop>GAL80* and *ey-Flp* (Fig. 2h). To restrict *Ten-a* expression to central neurons, we expressed *ten-a* RNAi in all ORNs (Fig. 2i). We found that *Ten-a* was highly expressed in a subset of ORNs and central neurons, and also showed a matching expression in the five glomeruli we focused on (Fig. 2h, i). The glomerular-specific differential *Ten-a* expression in central neurons probably arises mainly from PNs as they target dendrites to specific glomeruli, and punctate *Ten-a* staining was observed in PN cell bodies (Supplementary Fig. 5). In summary, *Ten-m* and *Ten-a* are each highly expressed in a distinct, but partially overlapping, subset of matching ORNs and PNs (Fig. 2j).

Teneurins are required for PN-ORN matching

To examine whether Teneurins are required for proper PN-ORN matching, we performed tissue-specific RNAi (Fig. 3 and Supplementary Fig. 2c) in all neurons using *C155-GAL4*, in PNs using *GH146-GAL4*, or in ORNs using *peb-GAL4*. To label specific subsets of PN dendrites independent of *GAL4*-UAS, we used the Q binary expression system³⁰, and converted *Mz19-GAL4* to *Mz19-QF* by bacterial artificial chromosome (BAC) recombineering (Supplementary Fig. 2d). We could thus perform *GAL4*-based RNAi knockdown while labelling PN dendrites and ORN axons in two colours independent of *GAL4*. We focused our analysis on Mz19 dendrites and Or47b axons, which innervate neighbouring glomeruli but never intermingle in wild type (Figs 1b and 3a, b).

Pan-neuronal RNAi of both *teneurin* genes shifted Or47b axons to a position between two adjacent Mz19 glomeruli, DA1 and VA1d (Fig. 3c). Moreover, Mz19 dendrites and Or47b axons intermingled without a clear border (Fig. 3c, d), reflecting a PN-ORN matching defect. We confirmed this using independent RNAi lines targeting

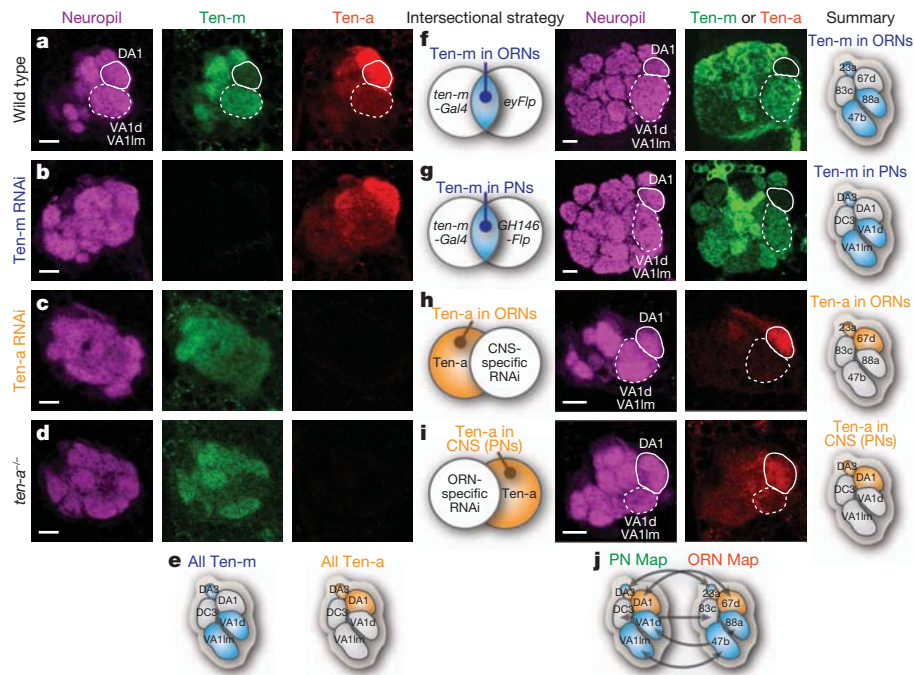


Figure 2 | Ten-m and Ten-a are differentially expressed in matching PN and ORN classes. **a**, Developing antennal lobes at 48 h APF stained by antibodies against Ten-m, Ten-a, and a neuropil marker, N-cadherin. Solid lines encircle the DA1 glomerulus (Ten-m low, Ten-a high). Dashed lines encircle the VA1d/VA1lm glomeruli (Ten-m high, Ten-a low). **b**, **c**, Ten-m and Ten-a proteins are undetectable following pan-neuronal RNAi of *ten-m* (**b**) and *ten-a* (**c**), respectively. **d**, A *ten-a* homozygous mutant eliminated the Ten-a antibody staining. **e**, Summary of elevated Ten-m and Ten-a expression in five select

glomeruli. **f**, **g**, Expression of the Flp-out GFP reporter *UAS>stop>mCD8GFP* at the intersection of *ten-m-GAL4* with ORN-specific *ey-Flp* (**f**) or with PN-specific *GH146-Flp* (**g**) in adult. **h**, **i**, Antibody staining of Ten-a in central nervous system (CNS)-neuron-specific RNAi (**h**) or in ORN-specific RNAi (**i**) at 48 h APF. **f**–**i**, Right, individual cell-type-specific Teneurin expression patterns are schematically summarized. **j**, Combined expression patterns of Teneurin proteins in PNs (left) and ORNs (right). Blue, Ten-m high; orange, Ten-a high. Scale bars, 10 μ m.

different regions of the *ten-m* and *ten-a* transcripts (Supplementary Fig. 6). Further, knocking down *ten-m* and *ten-a* only in PNs or only in ORNs also led to Mz19–Or47b intermingling (Fig. 3e and Supplementary Fig. 7a, d), indicating that Teneurins are required in both PNs and ORNs to ensure proper matching.

Next, we examined the contribution of each Teneurin by individual RNAi knockdown in ORNs. Knocking down *ten-m* and, to a lesser extent, *ten-a*, caused mild mismatching (Fig. 3e and Supplementary Fig. 7). This was greatly enhanced by simultaneous knockdown of both *ten-m* and *ten-a* (Fig. 3e), probably because Mz19–Or47b

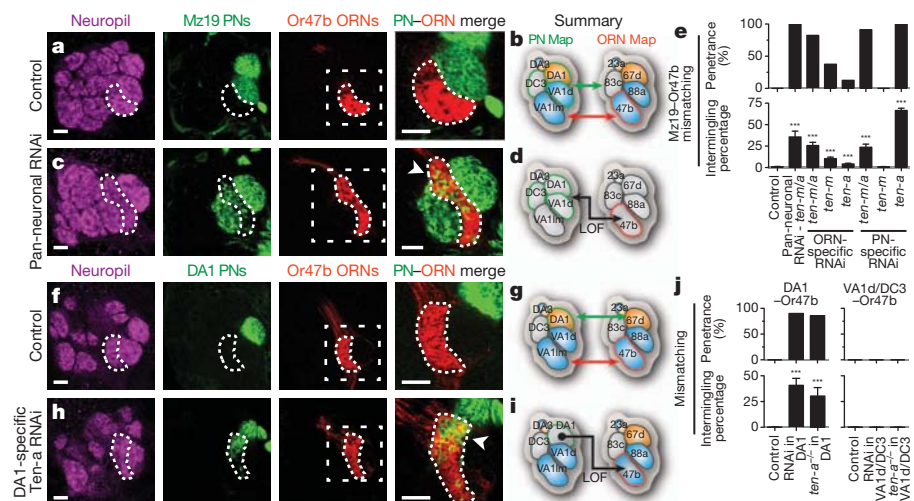


Figure 3 | Loss of Teneurins causes PN–ORN mismatching. **a**, Normally, Mz19 dendrites (green) innervate glomeruli adjacent to the VA1lm glomerulus, which is itself innervated by Or47b axons (red). The dashed line encircles Or47b axons. DC3 PNs are located posterior to DA1/VA1d PNs and Or47b ORNs, and are not visible in these sections. **c**, Mismatching phenotypes in *ten-m* and *ten-a* RNAi driven by the pan-neuronal driver *C155-GAL4*. Dashed lines encircle Or47b ORN axons, showing intermingling with Mz19 PN dendrites (arrowhead). **e**, Quantification of Mz19–Or47b mismatching phenotypes. For

all genotypes, $n \geq 15$. **f**, In control, DA1 PNs do not intermingle with Or47b ORNs. **h**, MARCM expression of *ten-a* RNAi in DA1 PNs causes dendrite intermingling with Or47b axons (arrowhead). **j**, Quantification of mismatching phenotypes. For all genotypes, $n \geq 6$. Error bars represent s.e.m. ***, $P < 0.001$ compared to control. **b**, **d**, **g**, **i**, Summary showing normal connectivity in control (**a**, **f**) and mismatching phenotypes following *teneurin* RNAi (**c**, **h**). Blue, Ten-m high; orange, Ten-a high. Green outlines, labelled PNs. Red outlines, labelled ORNs. Scale bars, 10 μ m.

mismatching requires weakening connections with their respective endogenous partners (Supplementary Fig. 7g). This synergy implies that multiple matching molecules can enhance partner matching robustness.

We also tested the functions of individual Teneurins in PNs. We found that the Mz19–Or47b mismatching was caused by PN-specific knockdown of *ten-a*, but not *ten-m* (Fig. 3e and Supplementary Fig. 7). As VA1d/DC3 and DA1 PNs arise from separate neuroblast lineages³¹, we used mosaic analysis with a repressible cell marker (MARCM) to generate neuroblast clones to label and knockdown *ten-a* in DA1 or VA1d/DC3 PNs (Fig. 3f–j; see Methods). *ten-a* knockdown only in DA1 PNs (normally Ten-a high) caused their dendrites to mismatch with Or47b axons (Fig. 3h–j). By contrast, *ten-a* knockdown in VA1d/DC3 PNs (normally Ten-a low) did not cause mismatching (Fig. 3j and Supplementary Fig. 8a, b). Similarly, MARCM loss-of-function of *ten-a* mutant in DA1 but not in VA1d/DC3 PNs resulted in mismatching with Or47b ORNs (Fig. 3j and Supplementary Fig. 8c, d). Thus, removal of *ten-a* from Ten-a-high DA1 PNs caused their dendrites to mismatch with Ten-a-low Or47b ORNs (Fig. 3i). The differential requirements of Ten-m and Ten-a in ORNs or PNs in preventing Mz19–Or47b mismatching probably reflect differential expression of Ten-m and Ten-a in the mismatching partners.

Our finding that loss of *ten-a* caused Ten-a-high PNs to mismatch with Ten-a-low ORNs (Fig. 3i, j), together with the matching expression of Teneurins proteins in PNs and ORNs, raised the possibility that Teneurins instruct class-specific PN–ORN connections through homophilic attraction: PNs expressing high-level Ten-m or Ten-a connect to ORNs with high-level Ten-m or Ten-a, respectively.

Teneurins instruct matching specificity

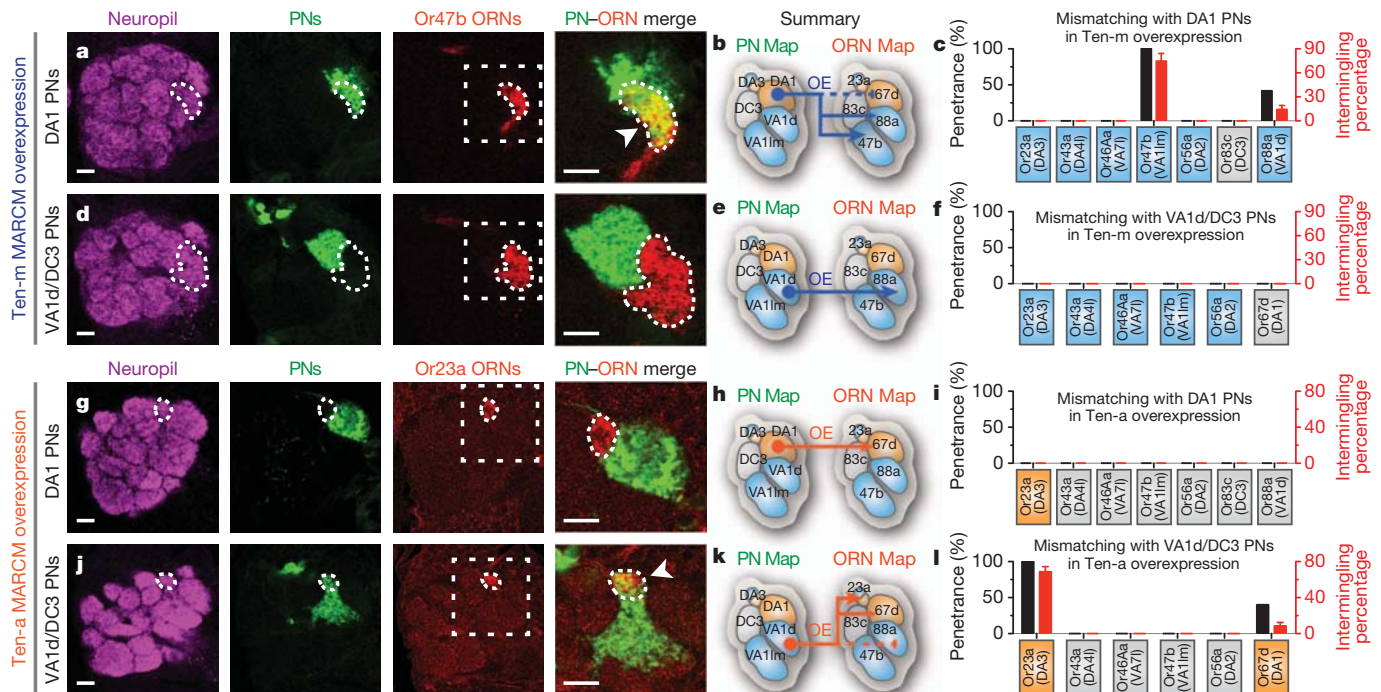
This homophilic attraction hypothesis predicts that overexpression of a given Teneurin in PNs (1) should preferentially affect PNs normally

expressing low levels of that Teneurin, causing their dendrites to lose endogenous connections with their cognate ORNs, and (2) should cause these PNs to make ectopic connections with ORNs expressing high levels of that Teneurin.

To test the first prediction, we examined whether Teneurin overexpression in Mz19 PNs impaired their endogenous connections with cognate ORNs. Consistent with our prediction, Ten-m overexpression specifically disrupted the connections of DA1 PNs and Or67d ORNs, a PN–ORN pair expressing low-level Ten-m (Supplementary Fig. 9b, e). Connections of the other two pairs were unaffected (Supplementary Fig. 9a, c, d, f). Likewise, Ten-a overexpression specifically disrupted connections between VA1d PNs and Or88a ORNs, a PN–ORN pair expressing low-level Ten-a (Supplementary Fig. 9g), but not between the other two PN–ORN pairs (Supplementary Fig. 9h, i).

To test the second prediction, we examined the specificity of ectopic connections made by Mz19 PNs overexpressing Teneurins, and sampled five non-partner ORN classes that project axons to the vicinity of Mz19 dendrites (Supplementary Fig. 10). We found that Ten-m overexpression in Mz19 PNs caused their dendrites to mismatch only with Or47b ORNs (Supplementary Fig. 10f). To examine additional mismatching phenotypes that may occur within Mz19 glomeruli and to determine whether DA1 or VA1d/DC3 PNs contribute to the ectopic connections, we used MARCM to overexpress Ten-m in individual PN classes. We found that Ten-m overexpression in DA1 PNs (Ten-m low) caused their dendrites to mismatch with Or47b (Fig. 4a, b) and (to a lesser extent) Or88a ORNs (Fig. 4b, c), both endogenously expressing high-level Ten-m. By contrast, Ten-m overexpression in VA1d/DC3 PNs did not produce ectopic connections with any non-matching ORNs tested (Fig. 4d–f).

Likewise, Ten-a overexpression in Mz19 PNs caused their dendrites to mismatch only with Or23a ORNs among all non-matching ORN classes sampled outside the Mz19 region (Supplementary Fig. 10l).



DA1 PNs (**g**), causes their dendrites to mismatch with Or23a axons. *P/GS/9267* and *P/GE/1914* (Supplementary Fig. 2) are used to overexpress Ten-m and Ten-a, respectively. **c, f, i, l**, Quantification of mismatching phenotypes ($n = 9$ for each). Error bars represent s.e.m. See Supplementary Fig. 10 for details on some genotypes quantified here. **b, e, h, k**, Schematic summarizing the mismatching phenotypes in Fig. 4 and Supplementary Figs 9 and 10. Blue, Ten-m high; orange, Ten-a high. Scale bars, 10 μ m.

Further, MARCM overexpression of Ten-a in VA1d/DC3 PNs (Ten-a low) caused their dendrites to mismatch specifically with Or23a (Fig. 4j, k) and (to a lesser extent) Or67d ORNs (Fig. 4k, l), both endogenously expressing high-level Ten-a (Fig. 4l). By contrast, Ten-a overexpression in DA1 PNs (Ten-a high) did not produce ectopic connections with any non-matching ORNs tested (Fig. 4g–i). Thus, both Ten-m and Ten-a overexpression analyses support the homophilic attraction hypothesis.

Our data also suggest that additional molecule(s) are required to determine completely the wiring specificity of the five PN–ORN pairs examined. For example, VA1d–Or88a and VA1m–Or47b have indistinguishable Ten-m/Ten-a expression patterns (Fig. 2j), and may require additional molecules to distinguish target choice. Indeed, Ten-a knockdown (Fig. 3h–j and Supplementary Fig. 8e, f) or Ten-m overexpression (Fig. 4b, c) caused DA1 PNs to mismatch preferentially with Or47b as opposed to Or88a axons. This suggests that the non-adjacent DA1 and VA1m share a more similar Teneurin-independent cell-surface code than the adjacent VA1d and VA1m. Likewise, Ten-a overexpression caused VA1d PNs to mismatch with the non-adjacent Or23a more so than the adjacent Or67d ORNs, even though both ORNs express high-level Ten-a (Fig. 4k, l). Finally, Ten-m overexpression in DC3 PNs, which express low-level Ten-m, did not change its matching specificity (Fig. 4f and Supplementary Fig. 9f), suggesting that Teneurin-independent mechanisms are involved in matching DC3 PNs and Or83c ORNs.

In summary, we showed that Teneurin overexpression in Teneurin-low PNs caused their dendrites to lose endogenous connections with Teneurin-low ORNs and mismatch with Teneurin-high ORNs (Fig. 4b, k). However, Teneurin overexpression in Teneurin-high PNs did not disrupt their proper connections (Fig. 4e, h). These data indicate that Teneurins instruct connection specificity, probably through homophilic attraction, by matching Ten-m or Ten-a levels in PN and ORN partners.

Ten-m promotes PN–ORN homophilic attractions

To test whether Teneurins interact *in vitro*, we separately transfected two populations of *Drosophila* S2 cells with Flag- and haemagglutinin (HA)-tagged Teneurins, and performed co-immunoprecipitations from lysates of these cells after mixing. We detected strong homophilic interactions between Flag- and HA-tagged Ten-m proteins and, to a lesser extent, between Flag- and HA-tagged Ten-a proteins (Fig. 5a). Ten-m and Ten-a also exhibited heterophilic interactions (Fig. 5a), which may account for their role in synapse organization²⁶.

Next, we tested whether Teneurins can homophilically promote *in vivo* trans-cellular interactions between PN dendrites and ORN axons. We simultaneously overexpressed Ten-m in Mz19 PNs using *Mz19-QF*, and Or67a and Or49a ORNs using *AM29-GAL4* (ref. 32; Fig. 5b). This enabled us to label and manipulate independently Mz19 dendrites and AM29 axons with distinct markers and transgenes. We chose *AM29-GAL4* because of its early onset of expression, whereas other class-specific ORN drivers start to express only after PN–ORN connection is established^{5,6}. AM29 axons do not normally connect with Mz19 dendrites (Fig. 5c, d).

Simultaneous overexpression of Ten-m in both Mz19 PNs and AM29 ORNs produced ectopic connections between them (Fig. 5c, g), suggesting that Ten-m homophilically promotes PN–ORN attraction. By contrast, Ten-m overexpression only in PNs or ORNs did not produce any ectopic connections, despite causing dendrite or axon mistargeting, respectively (Fig. 5c, e, f). These data ruled out the involvement of heterophilic partners in Ten-m-mediated attraction. Simultaneous overexpression of Ten-a in Mz19 PNs and AM29 ORNs did not produce ectopic connections (data not shown), possibly due to lower expression or weaker Ten-a homophilic interactions (Fig. 5a). Although heterophilic interactions between Ten-m and Ten-a also occur *in vitro* (Fig. 5a), heterophilic overexpression of Ten-m and

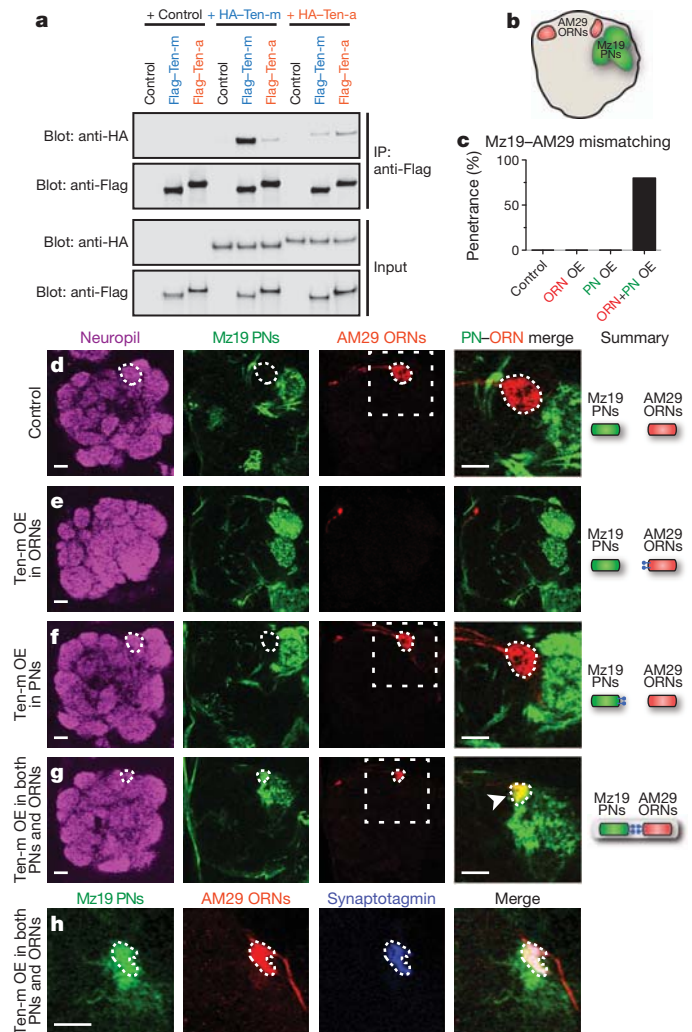


Figure 5 | Ten-m promotes homophilic interactions *in vitro* and *in vivo*.

a, Co-immunoprecipitation of Flag- and HA-tagged Teneurin proteins from separately transfected S2 cells. Co-immunoprecipitated HA-tagged Teneurin proteins are detected by anti-HA antibody and immunoprecipitated Flag-tagged Teneurin proteins by anti-Flag antibody (top two blots). The input lysates are immunoblotted for HA and Flag (bottom two blots). **b**, Schematic showing the relative positions of glomeruli targeted by Mz19 PN dendrites (green) and AM29 ORN axons (red). **c**, Quantification of mismatching between Mz19 PNs and AM29 ORNs ($n = 10$ for each condition). OE, overexpression. **d**, In control, Mz19 dendrites do not connect with AM29 axons. **e, f**, Overexpression of Ten-m only in AM29 ORNs (e) or Mz19 PNs (f) does not produce mismatching between them. Following Ten-m overexpression, AM29 axons mistarget posteriorly to Mz19 dendrites and are therefore not visible in e. **g**, Simultaneous overexpression of Ten-m in both PNs and ORNs produces ectopic Mz19–AM29 connections (arrowhead). Schematics on the right show the Mz19–AM29 connectivity in different conditions. **h**, The synaptic vesicle marker Synaptotagmin is enriched at these Mz19–AM29 ectopic connections. AM29 axons are labelled by *AM29-Gal4* with *UAS-mtdT* to visualize the entire axonal processes and *UAS-synaptotagmin-HA* to visualize synaptic vesicles in axon terminals. Mz19 dendrites are labelled by *Mz19-QF* driving *QUAS-mCD8GFP*. To overexpress Ten-m, *P{GS}9267* and *QUAS-ten-m* (Supplementary Fig. 2) are driven by *AM29-GAL4* and *Mz19-QF*, respectively. Scale bars, 10 μ m.

Ten-a in AM29 ORNs and Mz19 PNs did not produce ectopic connections (data not shown).

Finally, we examined whether these ectopic connections lead to the formation of synaptic structures. Indeed, the ectopic connections between Mz19 dendrites and AM29 axons were enriched in synaptotagmin–HA expressed from AM29 ORNs (Fig. 5h), suggesting that these connections can aggregate synaptic vesicles and could

be functional. We propose that Teneurins promote attraction between PN–ORN synaptic partners through homophilic interactions, eventually leading to synaptic connections.

Discussion

Compared to axon guidance, relatively little is known about synaptic target selection mechanisms^{2–4}. Among the notable examples, the graded expressions of vertebrate EphA and ephrin-A instruct the topographic targeting of retinal ganglion cell axons^{4,33–35}. Chick DSCAM and Sidekick promote lamina-specific arborization of retinal neurons³⁶. *Drosophila* Capricious promotes target specificity of photoreceptor and motor axons^{16,37–39}. *Caenorhabditis elegans* SYG-1 and SYG-2 specify synapse location through interaction between pre-synaptic axons and intermediate guidepost cells⁴⁰. However, it is unclear whether any of these molecules mediate direct, selective interactions between individual pre- and postsynaptic partners. Indeed, in complex neural circuits, it is not clear a priori whether molecular determinants mediate such interactions. For example, the final retinotopic map is thought to result from both ephrin signalling and spontaneous activity^{41,42}. Mammalian ORN axon targeting involves extensive axon–axon interactions through activity-dependent and independent modes^{43,44}, with minimal participation of post-synaptic neurons identified thus far.

Here, we show that Teneurins instruct PN–ORN matching through homophilic attraction. Although each glomerulus contains many synapses between cognate ORNs and PNs, these synapses transmit the same information and can be considered identical with regard to specificity. Thus, Teneurins represent a strong case in determining connection specificity directly between pre- and postsynaptic neurons. We further demonstrate that molecular determinants can instruct connection specificity of a moderately complex circuit at the level of individual synapses.

Our study reveals a requirement for PN–ORN attraction in the stepwise assembly of the olfactory circuit. PN dendrites and ORN axons first independently project to appropriate regions using global cues, dendrite–dendrite and axon–axon interactions^{8,9,12–14}. The initial, independent targeting of PN dendrites and ORN axons is eventually coordinated in their final one-to-one matching. We identified Teneurins as the first molecules to mediate this matching process, through direct PN–ORN attraction. Our analyses have focused on a subset of PN–ORN pairs involving trichoid ORNs⁴⁵, including Or67d, Or88a and Or47b that have been implicated in pheromone sensation⁴⁶. The partially overlapping expressions of Teneurins in other PN and ORN classes (Fig. 2 and Supplementary Fig. 4) suggest a broader involvement of Teneurins. At the same time, additional cell-surface molecules are also needed to determine completely connection specificity of all 50 PN–ORN pairs.

Teneurins are present throughout Animalia (Fig. 1h). Different vertebrate teneurins are broadly expressed in distinct and partially overlapping patterns in the nervous system¹⁸. Teneurin-3 is expressed in the visual system and is required for ipsilateral retinogeniculate projections⁴⁷. Our study suggests that differential Teneurin expression may have a general role in matching pre- and postsynaptic partners. Indeed, high-level Ten-m is involved in matching select motor neurons and muscles²⁶. Furthermore, Ten-m and Ten-a also trans-synaptically mediate neuromuscular synapse organization²⁶. This suggests that the synapse partner matching function of Teneurins may have evolved from their basal role in synapse organization. Interestingly, synaptic partner matching only involves homophilic interactions (this study and ref. 26), whereas synapse organization preferentially involves heterophilic interactions²⁶. This could not be fully accounted for by the different strength of their homophilic and heterophilic interactions *in vitro* (Fig. 5a). We speculate that these dual functions of Teneurins *in vivo* may engage signalling mechanisms that further distinguish homophilic versus heterophilic interactions.

METHODS SUMMARY

Detailed methods on fly stocks, generation of the *ten-a* allele, construction of transgenic flies, clonal analysis, histology, imaging, quantification and statistical analysis, epitope-tagged constructs, and co-immunoprecipitation can be found in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 June 2011; accepted 6 February 2012.

Published online 18 March 2012.

- Sperry, R. W. Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc. Natl Acad. Sci. USA* **50**, 703–710 (1963).
- Dickson, B. J. Molecular mechanisms of axon guidance. *Science* **298**, 1959–1964 (2002).
- Zipursky, S. L. & Sanes, J. R. Chemoaffinity revisited: dscams, protocadherins, and neural circuit assembly. *Cell* **143**, 343–353 (2010).
- Luo, L. & Flanagan, J. G. Development of continuous and discrete neural maps. *Neuron* **56**, 284–300 (2007).
- Komiyama, T. & Luo, L. Development of wiring specificity in the olfactory system. *Curr. Opin. Neurobiol.* **16**, 67–73 (2006).
- Brochtrup, A. & Hummel, T. Olfactory map formation in the *Drosophila* brain: genetic specificity and neuronal variability. *Curr. Opin. Neurobiol.* **21**, 85–92 (2011).
- Jefferis, G. S. X. E. *et al.* Developmental origin of wiring specificity in the olfactory system of *Drosophila*. *Development* **131**, 117–130 (2003).
- Komiyama, T., Sweeney, L. B., Schuldiner, O., Garcia, K. C. & Luo, L. Graded expression of semaphorin-1a cell-autonomously directs dendritic targeting of olfactory projection neurons. *Cell* **128**, 399–410 (2007).
- Hong, W. *et al.* Leucine-rich repeat transmembrane proteins instruct discrete dendrite targeting in an olfactory map. *Nature Neurosci.* **12**, 1542–1550 (2009).
- Hummel, T. *et al.* Axonal targeting of olfactory receptor neurons in *Drosophila* is controlled by Dscam. *Neuron* **37**, 221–231 (2004).
- Hummel, T. & Zipursky, S. L. Afferent induction of olfactory glomeruli requires N-cadherin. *Neuron* **42**, 77–88 (2004).
- Sweeney, L. B. *et al.* Temporal target restriction of olfactory receptor neurons by Semaphorin-1a/PlexinA-mediated axon-axon interactions. *Neuron* **53**, 185–200 (2007).
- Lattmann, M. *et al.* Semaphorin-1a controls receptor neuron-specific axonal convergence in the primary olfactory center of *Drosophila*. *Neuron* **53**, 169–184 (2007).
- Chou, Y.-H., Zheng, X., Beachy, P. A. & Luo, L. Patterning axon targeting of olfactory receptor neurons by coupled Hedgehog signaling at two distinct steps. *Cell* **142**, 954–966 (2010).
- Zhu, H. *et al.* Dendritic patterning by Dscam and synaptic partner matching in the *Drosophila* antennal lobe. *Nature Neurosci.* **9**, 349–355 (2006).
- Kurusu, M. *et al.* A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection. *Neuron* **59**, 972–985 (2008).
- Tucker, R. P. & Chiquet-Ehrismann, R. Teneurins: a conserved family of transmembrane proteins involved in intercellular signaling during development. *Dev. Biol.* **290**, 237–245 (2006).
- Tucker, R. P., Kenzelmann, D., Trzebiatowska, A. & Chiquet-Ehrismann, R. Teneurins: transmembrane proteins with fundamental roles in development. *Int. J. Biochem. Cell Biol.* **39**, 292–297 (2007).
- Young, T. R. & Leamey, C. A. Teneurins: important regulators of neural circuitry. *Int. J. Biochem. Cell Biol.* **41**, 990–993 (2009).
- Baumgartner, S. & Chiquet-Ehrismann, R. *Ten^a*, a *Drosophila* gene related to tenascin, shows selective transcript localization. *Mech. Dev.* **40**, 165–176 (1993).
- Baumgartner, S., Martin, D., Hagios, C. & Chiquet-Ehrismann, R. *Tenm*, a *Drosophila* gene related to tenascin, is a new pair-rule gene. *EMBO J.* **13**, 3728–3740 (1994).
- Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. *Nature Genet.* **43**, 977–983 (2011).
- Levine, A. *et al.* *odd Oz*: a novel *Drosophila* pair rule gene. *Cell* **77**, 587–598 (1994).
- Zheng, L. *et al.* *Drosophila* *Ten-m* and *Filamin* affect motor neuron growth cone guidance. *PLoS ONE* **6**, e22956 (2011).
- Liebl, F. L. W. *et al.* Genome-wide *P*-element screen for *Drosophila* synaptogenesis mutants. *J. Neurobiol.* **66**, 332–347 (2006).
- Mosca, T. J., Hong, W., Dani, V. S., Favaloro, V. & Luo, L. Trans-synaptic Teneurin signalling in neuromuscular synapse organization and target choice. *Nature* <http://dx.doi.org/10.1038/nature10923> (this issue).
- Li, H., Bishop, K. M. & O'Leary, D. D. M. Potential target genes of *EMX2* include *Odz/Ten-M* and other gene families with implications for cortical patterning. *Mol. Cell. Neurosci.* **33**, 136–149 (2006).
- Rubin, B. P., Tucker, R. P., Brown-Luedi, M., Martin, D. & Chiquet-Ehrismann, R. Teneurin 2 is expressed by the neurons of the thalamofugal visual system *in situ* and promotes homophilic cell-cell adhesion *in vitro*. *Development* **129**, 4697–4705 (2002).
- Oohashi, T. *et al.* Mouse *Ten-m/Odz* is a new family of dimeric type II transmembrane proteins expressed in many tissues. *J. Cell Biol.* **145**, 563–577 (1999).

30. Potter, C. J., Tasic, B., Russler, E. V., Liang, L. & Luo, L. The Q system: a repressible binary system for transgene expression, lineage tracing, and mosaic analysis. *Cell* **141**, 536–548 (2010).
31. Jefferis, G. S., Marin, E. C., Stocker, R. F. & Luo, L. Target neuron prespecification in the olfactory map of *Drosophila*. *Nature* **414**, 204–208 (2001).
32. Endo, K., Aoki, T., Yoda, Y., Kimura, K.-I. & Hama, C. Notch signal organizes the *Drosophila* olfactory circuitry by diversifying the sensory neuronal lineages. *Nature Neurosci.* **10**, 153–160 (2007).
33. Drescher, U. *et al.* *In vitro* guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. *Cell* **82**, 359–370 (1995).
34. Cheng, H. J., Nakamoto, M., Bergemann, A. D. & Flanagan, J. G. Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. *Cell* **82**, 371–381 (1995).
35. Feldheim, D. A. *et al.* Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron* **25**, 563–574 (2000).
36. Yamagata, M. & Sanes, J. R. Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. *Nature* **451**, 465–469 (2008).
37. Shinza-Kameda, M., Takasu, E., Sakurai, K., Hayashi, S. & Nose, A. Regulation of layer-specific targeting by reciprocal expression of a cell adhesion molecule, Capricious. *Neuron* **49**, 205–213 (2006).
38. Shishido, E., Takeichi, M. & Nose, A. *Drosophila* synapse formation: regulation by transmembrane protein with Leu-rich repeats, CAPRICIOUS. *Science* **280**, 2118–2121 (1998).
39. de Wit, J., Hong, W., Luo, L. & Ghosh, A. Role of leucine-rich repeat proteins in the development and function of neural circuits. *Annu. Rev. Cell Dev. Biol.* **27**, 697–729 (2011).
40. Shen, K., Fetter, R. D. & Bargmann, C. I. Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1. *Cell* **116**, 869–881 (2004).
41. McLaughlin, T., Torborg, C. L., Feller, M. B. & O'Leary, D. D. M. Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development. *Neuron* **40**, 1147–1160 (2003).
42. Pfeiffenberger, C., Yamada, J. & Feldheim, D. A. Ephrin-As and patterned retinal activity act together in the development of topographic maps in the primary visual system. *J. Neurosci.* **26**, 12873–12884 (2006).
43. Imai, T. *et al.* Pre-target axon sorting establishes the neural map topography. *Science* **325**, 585–590 (2009).
44. Serizawa, S. *et al.* A neuronal identity code for the odorant receptor-specific and activity-dependent axon sorting. *Cell* **127**, 1057–1069 (2006).
45. Couto, A., Alenius, M. & Dickson, B. J. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr. Biol.* **15**, 1535–1547 (2005).
46. van der Goes van Naters, W. & Carlson, J. R. Receptors and neurons for fly odors in *Drosophila*. *Curr. Biol.* **17**, 606–612 (2007).
47. Leamey, C. A. *et al.* Ten_m3 regulates eye-specific patterning in the mammalian visual pathway and is required for binocular vision. *PLoS Biol.* **5**, e241 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank V. Favaloro for advice on biochemistry and D. Luginbuhl for technical assistance; K. Zinn for the EP collection; H. Zhu for the initial contribution; R. Wides and S. Baumgartner for *teneurin* reagents; B. Zhang, Bloomington, Kyoto, Harvard and Vienna Stock Centers for fly stocks; BestGene for injection service; and K. Shen, T. Clandinin, D. Berns, V. Favaloro, X. Gao, S. Hippenmeyer, C. Liu, K. Miyamichi and X. Yu for critiques. Supported by a National Institutes of Health (NIH) grant (R01 DC-005982 to L.L.), and Epilepsy, Neonatology and Developmental Biology Training Grants (NIH 5T32 NS007280 and HD007249 to T.J.M.). L.L. is an investigator of the Howard Hughes Medical Institute.

Author Contributions W.H. designed and performed all experiments. T.J.M. assisted in some experiments. L.L. supervised the project. W.H. and L.L. wrote the manuscript with feedback from T.J.M.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.L. (lluo@stanford.edu).

METHODS

Fly stocks. *Mz19-GAL4* (ref. 7) was used to label PNs. *Or-rCD2* lines^{15,48} (*Or47b-rCD2* and *Or88a-rCD2*), *Or-mCD8GFP* lines⁴⁵ (*Or23a-mCD8GFP*, *Or43a-mCD8GFP*, *Or46Aa-mCD8GFP*, *Or56a-mCD8GFP* and *Or83c-mCD8GFP*), *Or67d-GAL4* (ref. 49) and *AM29-GAL4* (ref. 32) were used to label ORNs. *GH146-Flp* (ref. 9), *ey-Flp* (ref. 50), *UAS>stop>mCD8GFP* (ref. 9) and *tubP>stop>GAL80* (ref. 51) were used to perform the intersectional expression analysis. *P{GS}9267 (ten-m)* was generated by the *Drosophila* Gene Search Project (Metropolitan University)⁵² and *P{GE}1914 (ten-a)*¹⁶ was from the GenExel collection of EP lines generated by the Korean Advanced Institute of Science and Technology. Their ability to drive the overexpression of each respective Teneurin was verified by elevated antibody staining.

All RNAi lines targeting *ten-m* or *ten-a* from the Vienna *Drosophila* RNAi Center (*UAS-ten-m*^{RNAi-V51173} and *UAS-ten-a*^{RNAi-V32482}), the Bloomington *Drosophila* Stock Center (*UAS-ten-m*^{RNAi-JF03323} and *UAS-ten-a*^{RNAi-JF03375}), and the National Institute of Genetics Fly Stock Center (*UAS-ten-m*^{RNAi-5723R} and *UAS-ten-a*^{RNAi-2590R}) were collected. The efficiency of all RNAi lines was tested by pan-neuronal expression using *C155-GAL4* followed by Ten-m or Ten-a antibody staining. *UAS-ten-m*^{RNAi-V51173} and *UAS-ten-m*^{RNAi-JF03323} targeting *ten-m*, and *UAS-ten-a*^{RNAi-V32482} and *UAS-ten-a*^{RNAi-JF03375} targeting *ten-a* were able to eliminate respective antibody staining beyond detection. *UAS-ten-m*^{RNAi-V51173} and *UAS-ten-a*^{RNAi-V32482} were used in all the experiments except Supplementary Fig. 6, in which *UAS-ten-m*^{RNAi-JF03323} and *UAS-ten-a*^{RNAi-JF03375} were used to confirm the RNAi phenotypes. *UAS-Dcr2* (ref. 53) was used to enhance RNAi efficiency.

To identify *ten-m-GAL4*, we collected a group of GAL4 enhancer traps⁵⁴ located near the 5' end of the *ten-m* gene. Their expression patterns were determined using a membrane-tagged GFP reporter *UAS-mCD8GFP* gene from the *Drosophila* Genetic Resource Center. *NP6658-GAL4*, which recapitulated the glomerulus-specific Ten-m staining pattern, was identified and referred to as *ten-m-GAL4*.

Generation of the *ten-a* allele. A small deficiency of *ten-a* was generated by FRT-mediated excision. This deficiency allele is homozygous viable and contains a deletion between *P{XP}d07540* and *P{Bac}WHJf01428*, which flanks the entire *ten-a* genomic region (~140 kb) and four additional predicted genes (Supplementary Fig. 2a), and is referred to as *Df(X)ten-a*. The deletion was verified by both PCR and antibody staining against Ten-a. The loss-of-function phenotypes were due to the loss of *ten-a* rather than the four additional predicted genes, as they mimicked the RNAi phenotypes.

Construction of UAS and QUAS transgenic flies. Ten-m and Ten-a coding sequences were amplified from the cDNA constructs^{23,55,56}. One primer amplified from the start codon and added a CACC overhang for the TOPO reaction and a Kozak sequence. The other primer amplified to the stop codon. The PCR products were subcloned into *pENTR-D/TOPO* (Invitrogen). A 46-bp irrelevant fragment was found in the middle of the *ten-m* coding sequence in the original cDNA construct, and was removed by replacing a small region containing this fragment with the corresponding region in the *ten-m* genomic DNA. To make *UAS-ten-m*, *UAS-ten-a*, *QUAS-ten-m* and *QUAS-ten-a*, *pENTR-ten-m* and *pENTR-ten-a* were recombined into destination vector *pUAS-Gateway-attB*⁵⁷ and *pQUAS-Gateway-attB* using LR Clonase II (Invitrogen). The destination vector *pQUAS-Gateway-attB* was constructed by replacing the UAS site in *pUAS-Gateway-attB* with a QUAS site. All constructs were sequence verified. All the UAS and QUAS transgenes were integrated into both *attP24* and *86Fb* landing sites^{58,59} on second and third chromosomes, respectively. All transgenic flies were verified by PCR and overexpression followed by antibody staining. The UAS and QUAS transgenes inserted in the *86Fb* site were used in this paper.

BAC recombineering to construct *Mz19-QF*. A 110-kb BAC (#CH321-85L03) in the *attB-P[acman]-CmR* vector⁶⁰, which contains genomic DNA that covers the *Mz19-GAL4* enhancer trap insertion site, was collected from the BACPAC Resources Center. The QF coding sequence, with a P-element minimal promoter and an hsp70 polyA, was amplified using primers containing 50-bp arm sequences allowing site-specific recombination. The 5-kb PCR product was recombined into the 110-kb genomic BAC using bacterial BAC recombineering and was verified by sequencing. The 115-kb *Mz19-QF* BAC was further verified by digestion pattern analysis and used to produce transgenes at the VK37 landing site⁶¹ on the second chromosome by BestGene. The *Mz19-QF* transgenic flies were verified by PCR and the expression of reporters *QUAS-mCD8GFP* or *QUAS-mTdt3HA*.

Clonal analysis. To determine the contribution of individual PNs to the ectopic connections with ORNs, the MARCM method⁶² was applied. Briefly, heat-shock-induced Flp activity caused mitotic recombination of the FRT chromosome arm such that one of the daughter cells lost *GAL80*. This cell (and its progeny) can therefore be labelled by the *GALA-UAS* system. For generating neuroblast clones, flies were heat-shocked between 24–48 h after egg laying for 1 h at 37 °C. *Mz19-GAL4* labels VA1d

and DC3 from the anterodorsal neuroblast and DA1 from the lateral neuroblast⁷. By generating neuroblast clones at 24–48 h after egg laying, we used *MZ19-GAL4* to specifically label DA1 or VA1d/DC3 PNs and simultaneously express RNAi targeting *ten-a*, or overexpress Ten-m or Ten-a in the labelled neurons.

In the *ten-a* mutant analysis, *Df(X)ten-a* was placed in *trans* to *GAL80* on the FRT chromosome arm. Upon Flp-induced mitotic recombination, one of the daughter cells became homozygous for *ten-a* and simultaneously lost *GAL80*. We used *MZ19-GAL4* to specifically label DA1 or VA1d/DC3 mutant PNs.

Different classes of ORNs, except for Or67d, were labelled by *Or-mCD8GFP* transgenes in a GAL4-independent manner, allowing the visualization of the specific matching between the labelled PNs and ORNs. Owing to the lack of an Or67d-*mCD8GFP*, Or67d ORNs were labelled by *Or67d-GAL4* and Ten-m overexpression was achieved by using *Mz19-QF* to drive *QUAS-ten-m* (Supplementary Fig. 9). In Teneurin overexpression by *Mz19-QF*, Or67d-*GAL4* expression was found unchanged compared with the control, and co-localized with Ncad staining in the DA1 glomerulus, which can be unambiguously identified (Supplementary Fig. 9). Therefore, Ncad staining in the DA1 glomerulus was used to determine the location of Or67d ORNs in Fig. 4f, l, in which Teneurins were overexpressed by *Mz19-GAL4*.

Histology. The procedures used for fixation and immunostaining were described previously⁶³. For primary antibodies, we used mouse nc82 (1:30), rat antibody to N-cadherin (1:40), rat antibody to mCD8 (1:100), mouse antibody to rCD2 (1:200), chicken antibody to GFP (1:1,000), mouse antibody to HA (1:1,000), rabbit antibody to HA (1:1,000), rabbit antibody to DsRed (1:500), mouse antibody (mAb20) to Ten-m (1:3,000)²³, and guinea pig antibody to Ten-a (1:100)⁶⁴. Neuropil staining indicates the antennal lobe, where PN dendrites and ORN axons are located. Fluorescent labelling outside the antennal lobe may come from labelled PN cell bodies or non-specific tissues.

Imaging, quantification and statistical analysis. Immunostained brains were imaged with a Zeiss LSM 510 Meta laser-scanning confocal microscope. Images of antennal lobes were taken as confocal stacks with 1-μm-thick sections. Representative single sections were shown to illustrate the matching and mismatching between PN dendrites and ORN axons. Penetration of phenotypes represents the percentage of animals in which at least one antennal lobe showed a given phenotype among the total animals examined. Percentage of intermingling represents the fraction of labelled dendrites located within the axonal area of a given ORN class, and was measured by dividing dendritic area by total axonal area in a single confocal plane that shows maximum intermingling between dendrites and axons. Statistical significance between two samples was determined by the unpaired Student's *t*-test.

Flag- and HA-tagged constructs. To express Flag- and HA-tagged proteins in S2 cells, the Gateway destination vectors *pUAS-Flag-Gateway(-w)* and *pUAS-HA-Gateway(-w)* were generated by removing a ~4.5-kb non-essential fragment between two DraIII sites that contains the *white* gene from the original Gateway vectors pTFW and pTHW (*Drosophila* gateway collection, DGRC, Bloomington), respectively. The modified destination vectors are ~40% smaller than the original ones while preserving all the essential components for S2 cell expression, and showed greater transfection and expression efficiency in S2 cells. To express Flag- and HA-tagged Teneurin proteins in S2 cells, *pENTR-ten-m* and *pENTR-ten-a* were recombined into modified destination vectors *pUAS-Flag-Gateway(-w)* and *pUAS-HA-Gateway(-w)* using LR Clonase II (Invitrogen). All expression constructs, including *UAS-Flag-ten-m*, *UAS-Flag-ten-a*, *UAS-HA-ten-m* and *UAS-HA-ten-a*, were sequence verified.

Co-immunoprecipitation assay. S2 cells were cultured in Schneider's insect medium (Sigma) according to the manufacturer's description. *UAS-Flag-ten-m*, *UAS-Flag-ten-a*, *UAS-HA-ten-m* or *UAS-HA-ten-a* constructs were separately transfected into S2 cells, along with an *Actin5c-GAL4* vector, using the Effectene transfection reagent (QIAGEN). The amount of each construct and the number of cells used for transfection were adjusted to ensure comparable expression levels of Ten-m and Ten-a proteins. Three days after transfection, separately transfected cells were harvested, mixed together, and incubated for 1 h at room temperature (25 °C). Equivalent amounts of untransfected cells were used as controls, and the final mixtures contained the same total amount of cells under all co-immunoprecipitation conditions. The mixed cells were lysed in lysis buffer (50 mM Tris-HCl pH 7.4, 10 mM MgCl₂, 150 mM NaCl, 1 mM EGTA, 10% glycerol) supplemented with 0.5% Nonidet P-40 and protease inhibitor cocktail (Sigma). The cell lysates were then incubated with EZview Red anti-Flag M2 affinity gel (Sigma) for 3 h at 4 °C with rotation. The samples were washed extensively in lysis buffer. The proteins were eluted in 2% SDS elution buffer, and were detected using western blot analysis using rat antibody to HA (1:1,000, Roche), mouse antibody to Flag (1:5,000, Sigma), and HRP-conjugated-goat antibodies to rat or mouse primaries (both at 1:20,000, Jackson ImmunoResearch).

48. Zhu, H. & Luo, L. Diverse functions of N-cadherin in dendritic and axonal terminal arborization of olfactory projection neurons. *Neuron* **42**, 63–75 (2004).

49. Kurtovic, A., Widmer, A. & Dickson, B. J. A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. *Nature* **446**, 542–546 (2007).
50. Newsome, T. P., Asling, B. & Dickson, B. J. Analysis of *Drosophila* photoreceptor axon guidance in eye-specific mosaics. *Development* **127**, 851–860 (2000).
51. Bohm, R. A. *et al.* A genetic mosaic approach for neural circuit mapping in *Drosophila*. *Proc. Natl Acad. Sci. USA* **107**, 16378–16383 (2010).
52. Toba, G. *et al.* The gene search system. A method for efficient detection and rapid molecular identification of genes in *Drosophila melanogaster*. *Genetics* **151**, 725–737 (1999).
53. Dietzl, G. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151–156 (2007).
54. Hayashi, S. *et al.* GETDB, a database compiling expression patterns and molecular locations of a collection of Gal4 enhancer traps. *Genesis* **34**, 58–61 (2002).
55. Fascetti, N. & Baumgartner, S. Expression of *Drosophila* Ten-a, a dimeric receptor during embryonic development. *Mech. Dev.* **114**, 197–200 (2002).
56. Kinel-Tahan, Y., Weiss, H., Dgany, O., Levine, A. & Wides, R. *Drosophila* *odz* gene is required for multiple cell types in the compound retina. *Dev. Dyn.* **236**, 2541–2554 (2007).
57. Tea, J. S., Chihara, T. & Luo, L. Histone deacetylase Rpd3 regulates olfactory projection neuron dendrite targeting via the transcription factor Prospero. *J. Neurosci.* **30**, 9939–9946 (2010).
58. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genet.* **40**, 476 (2008).
59. Bischof, J., Maeda, R. K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific ϕ C31 integrases. *Proc. Natl Acad. Sci. USA* **104**, 3312–3317 (2007).
60. Venken, K. J. T. *et al.* Versatile P[acman] BAC libraries for transgenesis studies in *Drosophila melanogaster*. *Nature Methods* **6**, 431–434 (2009).
61. Venken, K. J. T., He, Y., Hoskins, R. A. & Bellen, H. J. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**, 1747–1751 (2006).
62. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
63. Wu, J. S. & Luo, L. A protocol for dissecting *Drosophila melanogaster* brains for live imaging or immunostaining. *Nature Protocols* **1**, 2110–2115 (2006).
64. Rakovitsky, N. *et al.* *Drosophila* *Ten-a* is a maternal pair-rule and patterning gene. *Mech. Dev.* **124**, 911–924 (2007).

Structure of the mitotic checkpoint complex

William C. H. Chao^{1*}, Kiran Kulkarni^{1*}, Ziguang Zhang¹, Eric H. Kong¹ & David Barford¹

In mitosis, the spindle assembly checkpoint (SAC) ensures genome stability by delaying chromosome segregation until all sister chromatids have achieved bipolar attachment to the mitotic spindle. The SAC is imposed by the mitotic checkpoint complex (MCC), whose assembly is catalysed by unattached chromosomes and which binds and inhibits the anaphase-promoting complex/cyclosome (APC/C), the E3 ubiquitin ligase that initiates chromosome segregation. Here, using the crystal structure of *Schizosaccharomyces pombe* MCC (a complex of mitotic spindle assembly checkpoint proteins Mad2, Mad3 and APC/C co-activator protein Cdc20), we reveal the molecular basis of MCC-mediated APC/C inhibition and the regulation of MCC assembly. The MCC inhibits the APC/C by obstructing degron recognition sites on Cdc20 (the substrate recruitment subunit of the APC/C) and displacing Cdc20 to disrupt formation of a bipartite D-box receptor with the APC/C subunit Apc10. Mad2, in the closed conformation (C-Mad2), stabilizes the complex by optimally positioning the Mad3 KEN-box degron to bind Cdc20. Mad3 and p31^{comet} (also known as MAD2L1-binding protein) compete for the same C-Mad2 interface, which explains how p31^{comet} disrupts MCC assembly to antagonize the SAC. This study shows how APC/C inhibition is coupled to degron recognition by co-activators.

The fidelity of chromosome separation in mitosis is governed by an evolutionarily conserved cell-cycle checkpoint mechanism called the SAC^{1,2}. The SAC arrests the mitotically dividing cell to allow complete chromosome attachment to the bipolar mitotic spindle. The essence of the SAC is to block the onset of anaphase by inhibiting APC/C-mediated ubiquitin-dependent degradation of securin and mitotic cyclin. Components of the SAC that are responsible for detecting unattached kinetochores and for propagating signals to the APC/C have been identified^{3,4}, but the molecular basis underlying these processes is only partially understood^{1,2}. Mad2 and Mad3 (BubR1 in metazoans) mediate APC/C inhibition through their association with its co-activator subunit Cdc20 (refs 5–12). Mad2, Mad3 and Cdc20 (together with mitotic checkpoint protein Bub3) form the MCC that directly binds the APC/C to inhibit substrate recognition^{13,14}. Mad2 and Mad3 cooperate to antagonize Cdc20-dependent activation of the APC/C¹², with Mad3–Cdc20 interactions requiring the pre-assembly of a Cdc20–Mad2 complex^{12,15–18}. Thus, SAC signalling occurs through the generation of the Cdc20–Mad2 complex, a process initiated by Mad1, which is the Mad2 receptor at unattached kinetochores. Central to the association of Mad2 with Cdc20 is the inter-conversion of Mad2 between the open (O-Mad2) and closed (C-Mad2) structural states^{1,19}. These states of Mad2 differ in the topology of a carboxy-terminal β -sheet that repositions in C-Mad2 to enable binding to its protein ligands, Mad1 or Cdc20 (refs 20–22). In the template model for SAC activation, Mad1 interacts with C-Mad2, generating the C-Mad2–Mad1 complex that subsequently recruits O-Mad2 through the C-Mad2 dimerization interface. By inducing the conformational transition of O-Mad2 to C-Mad2, the Mad1-bound C-Mad2 subunit catalyses the binding of Mad2 to Cdc20 (refs 17, 23).

APC/C activity and its substrate recruitment are dependent on its co-activators (Cdc20 and Cdh1)²⁴, which recognize APC/C substrates through two destruction motifs (degrons); the D box²⁵ and the KEN (Lys–Glu–Asn) box²⁶. Mad3 contains a KEN box that is essential for MCC assembly^{15,27,28}, suggesting that Mad3 may act as a pseudosubstrate

to block substrate recognition by APC/C^{Cdc20}. However, other studies showing that the promotion of ubiquitin-mediated degradation of Cdc20 by the SAC is dependent on the KEN box of Mad3 (refs 18, 27, 29) indicate that there is a more complex mechanism controlling APC/C^{Cdc20} activity. The mechanisms underlying APC/C activation after SAC silencing are also poorly understood. In metazoans, p31^{comet} antagonizes the SAC³⁰ by functioning as a structural mimic of Mad2 that binds at the Mad2 dimerization interface to inhibit the conformational activation of O-Mad2 (ref. 31). Ubch10, assisted by p31^{comet}, catalyses Cdc20 ubiquitination, which leads to the disassembly of the MCC³².

To understand the molecular mechanisms underlying the mitotic checkpoint complex, we determined the crystal structure of the fission yeast MCC. The structure shows how Mad2 and Mad3 cooperatively inhibit Cdc20, and indicates how p31^{comet} would antagonize MCC assembly. The structure of Cdc20 in the context of the MCC offers the first opportunity to visualize degron recognition by co-activators. The interaction between Mad2 and Mad3 positions the Mad3 KEN box towards the KEN-box receptor of the Cdc20 WD40 domain. Additionally, an unexpected D-box mimic of the Mad3 C terminus reveals the D-box-binding site on Cdc20, thus demonstrating the structural basis of D-box recognition by co-activators.

Overall structure of the MCC

We generated the fission yeast MCC by co-expressing Cdc20, Mad2 and Mad3 in insect cells. The complex comprises Cdc20 with all functional domains (C box, Mad2-binding motif, WD40 domain and Ile–Arg tail (Supplementary Fig. 1)), Mad2 locked in its closed conformation that facilitates binding to Cdc20 (ref. 31), and Mad3 truncated after the tetratricopeptide repeat (TPR) domain³³ and thus lacking its C-terminal KEN box. Bub3 was omitted from the complex because previous studies indicated that Bub3 was not an integral part of MCC in fission yeast²⁸ and was not required for MCC-mediated inhibition of human APC/C³⁴.

¹Division of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London, SW3 6JB, UK.

*These authors contributed equally to this work.

The crystal structure of the MCC, at 2.3 Å resolution, shows that Cdc20, Mad2 and Mad3 assemble into a triangular heterotrimer (Fig. 1 and Supplementary Table 1). Mad3 coordinates the overall organization of the complex by forming numerous inter-subunit interactions with Mad2 and Cdc20, whereas Cdc20 and Mad2 interact primarily through the sequestering of the Mad2-binding motif of Cdc20 by the Mad2 'safety belt' (refs 21, 22). The core architecture of Mad3 is a contiguous TPR superhelix with three TPR motifs flanked by capping α -helices that resembles the BubR1 amino terminus³³. In addition, not present in the BubR1 structure, the conserved N terminus of Mad3 that incorporates the N-terminal KEN box that is essential for MCC assembly^{15,27,28} adopts a helix-loop-helix (HLH) motif (Fig. 1 and Supplementary Fig. 2). The HLH motif simultaneously binds Mad2 and Cdc20, orienting the KEN box towards its receptor on Cdc20 (Fig. 1a, b). Mad3 also contacts Mad2 and Cdc20 through its TPR domain (Fig. 1a, c).

The WD40 domain of Cdc20 conforms to a canonical seven-bladed β -propeller (Fig. 1a). The Mad2-binding motif, N-terminal to the WD40 domain, is structurally well defined, engaging the safety belt^{21,22} of Mad2 and adopting a conformation that is similar to the Mad2-binding motif of Mad1 bound to Mad2 (ref. 22) (Fig. 2a). The linker connecting the Mad2-binding motif with the N terminus of the WD40 domain is disordered, as are the two APC/C interacting motifs: a conserved C box that is immediately N-terminal to the Mad2-binding motif³⁵, and the C-terminal Ile-Arg tail that interacts with the TPR subunit Cdc27³⁶ (Supplementary Fig. 1).

Mad2 is an α/β -HORMA-class protein²⁰ (Fig. 1a and Supplementary Fig. 3). Its C-Mad2 conformation enables interactions with the

Mad2-binding motifs of Cdc20 and Mad1 and, as revealed in the MCC crystal structure, is the only state of Mad2 that can recognize Mad3. Mad2 interacts through its α -C helix and $\beta 8'$ – $\beta 8''$ hairpin with the HLH motif of Mad3 (Fig. 2b). Notably, such Mad2 interactions resemble those between C-Mad2 and p31^{comet} (ref. 31.) (Fig. 2c), and C-Mad2 in an asymmetric C-Mad2–O-Mad2 dimer³⁷ (Fig. 2d). The involvement of the $\beta 8'$ – $\beta 8''$ hairpin—a region of Mad2 that undergoes substantial conformational change upon transition from the open to closed states²¹—at the Mad2–Mad3 interface indicates that Mad3 binds exclusively to C-Mad2.

Mutations of the Mad2 α -C helix disrupt C-Mad2–Mad3 interactions³⁸, consistent with our structure. To test the Mad3 interface, we replaced Met 13 with Arg (Fig. 2b). This mutation dissociated Mad3 from a Cdc20–C-Mad2 heterodimer when size-exclusion chromatography was performed (Supplementary Fig. 4). Together, these data confirm the physiological relevance of the C-Mad2–Mad3 interface and show that C-Mad2 is required to confer high-affinity binding of Mad3 to Cdc20, consistent with studies showing that Mad3 association with Cdc20 *in vivo* is synergistic with Mad2 (refs 12, 15, 16).

The Cdc20 WD40 domain is a receptor for KEN and D box

We identified two highly conserved surfaces on the Cdc20 WD40 domain, and these are responsible for APC/C degron recognition: the KEN-box²⁶ receptor, situated on the top side of the WD40 domain at the centre of the β -propeller (Fig. 3a), and the D-box²⁵ co-receptor lying in a channel between blades 1 and 7 (Fig. 4a). The KEN-box residues of Mad3 (Lys 20, Glu 21 and Asn 22) emerge from the C

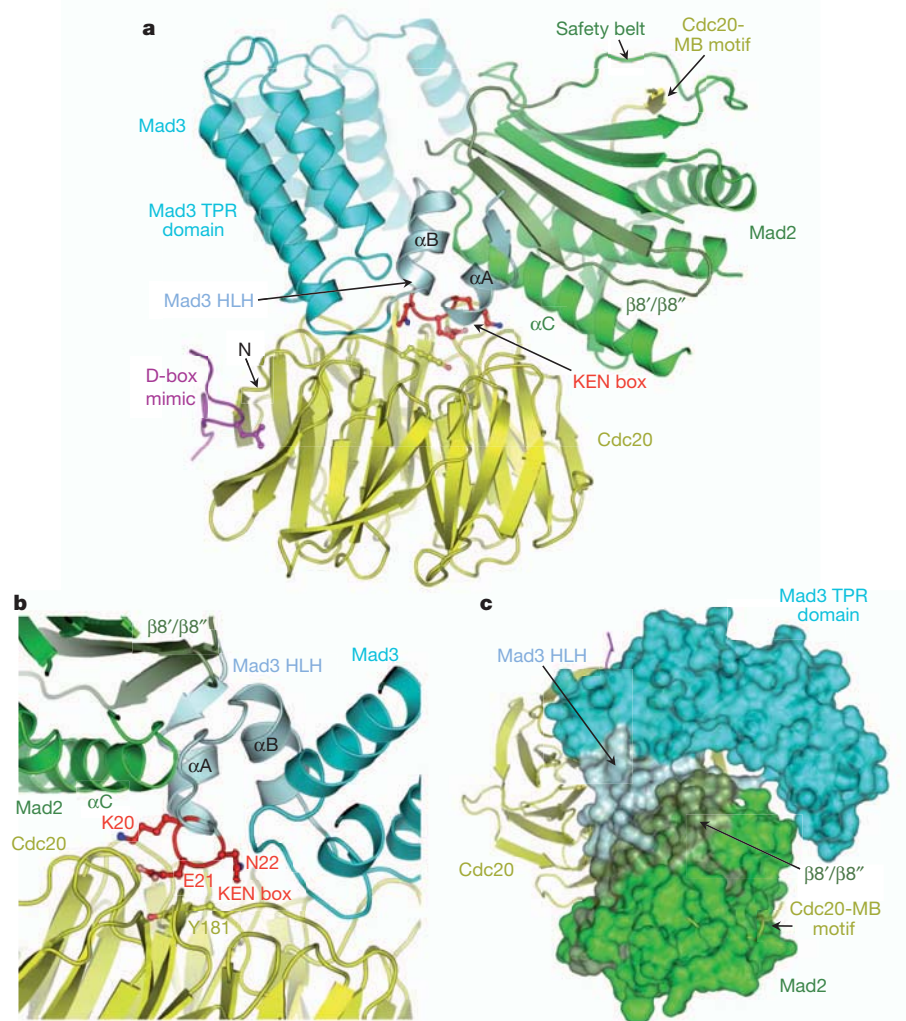


Figure 1 | Structure of *S. pombe* MCC trimer.

a, Cartoon representation of the complex of Mad2 (green), Mad3 (cyan) and Cdc20 (yellow). Mad2 is in the C-Mad2 conformation. The KEN box is shown in red, located in the HLH motif of Mad3. The D-box mimic (magenta) that is bound to Cdc20 is from the C terminus of Mad3 from a symmetry-related molecule. The Mad2 $\beta 8'$ – $\beta 8''$ hairpin (dark green) that forms the Mad2–Mad3 interface, repositions on conversion from O-Mad2 to C-Mad2. The N terminus of the WD40 domain is indicated. **b**, Details of the Mad3 HLH interaction with Mad2 and Cdc20. **c**, Surface representation of the MCC. The interaction between Mad2 and Mad3 positions the Mad3 KEN box at the KEN-box receptor, which is located at the centre of the top side of Cdc20's WD40 domain. MB, Mad2-binding.

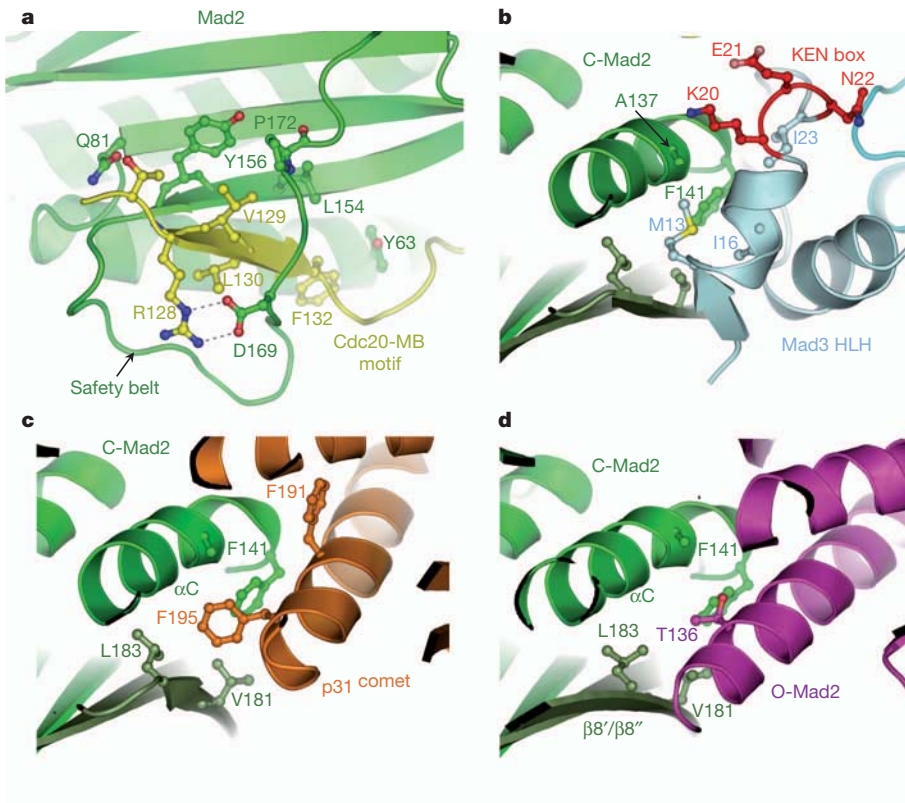


Figure 2 | Details of Cdc20–Mad2–Mad3 interactions. **a**, Mad2-binding motif of Cdc20 bound to the Mad2 safety belt. **b**, A universal Mad2 dimer interface contacts Mad3, p31^{comet} and O-Mad2. Mad2–Mad3 interactions. **c**, Mad2–p31^{comet} (ref. 31) (Protein Data Bank (PDB) accession 2QYF). **d**, C-Mad2–O-Mad2 (ref. 37) (PDB accession 2V64).

terminus of the α -A helix of the HLH as an underground turn of α -helix (Figs 1 and 3b). Their conformation resembles a 'U', with Glu 21 at the turn dipping into a depression at the centre of the Cdc20 β -propeller toroid. Tyr 181 of Cdc20, situated on an extended loop connecting blade 1 with blade 7, acts as a platform to support the aliphatic moiety of the Glu 21 side chain. Apart from this single hydrophobic interaction, the KEN box forms entirely polar contacts involving its side-chain and main-chain groups with conserved polar and charged residues of loops on four blades of the Cdc20 β -propeller. Notably, the five residues of Cdc20 that contribute side-chain interactions to the KEN box (Asp 180, Asn 326, Thr 368, Gln 392

and Arg 438) are invariant in both Cdc20 and Cdh1, indicating a universal mode of co-activator–KEN box interaction (Supplementary Fig. 5). A proline (Pro 25) positioned three residues C-terminal to the Mad3 KEN box—well conserved in KEN-box motifs—acts to break the α -helix and orient the polypeptide chain away from the Cdc20 surface.

The Mad3 KEN box has been proposed to function as a pseudo-substrate inhibitor, blocking access of Cdc20 to KEN-box degrons in APC/C substrates¹⁵. Thus, the mode of binding of the Mad3 KEN box to Cdc20 should serve as a model for understanding how co-activators recognize APC/C substrates. Consistent with this idea, mutating the

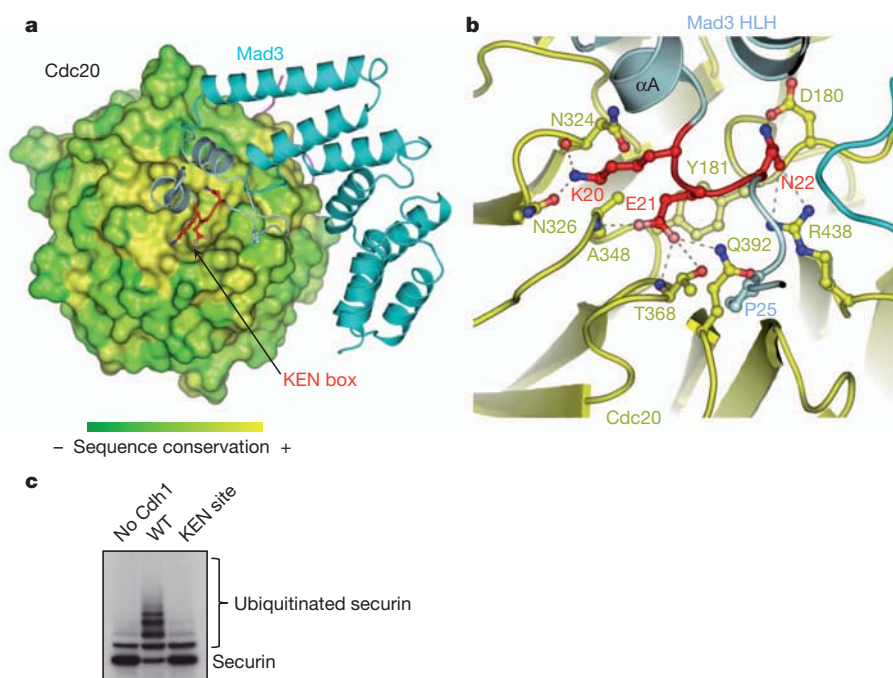
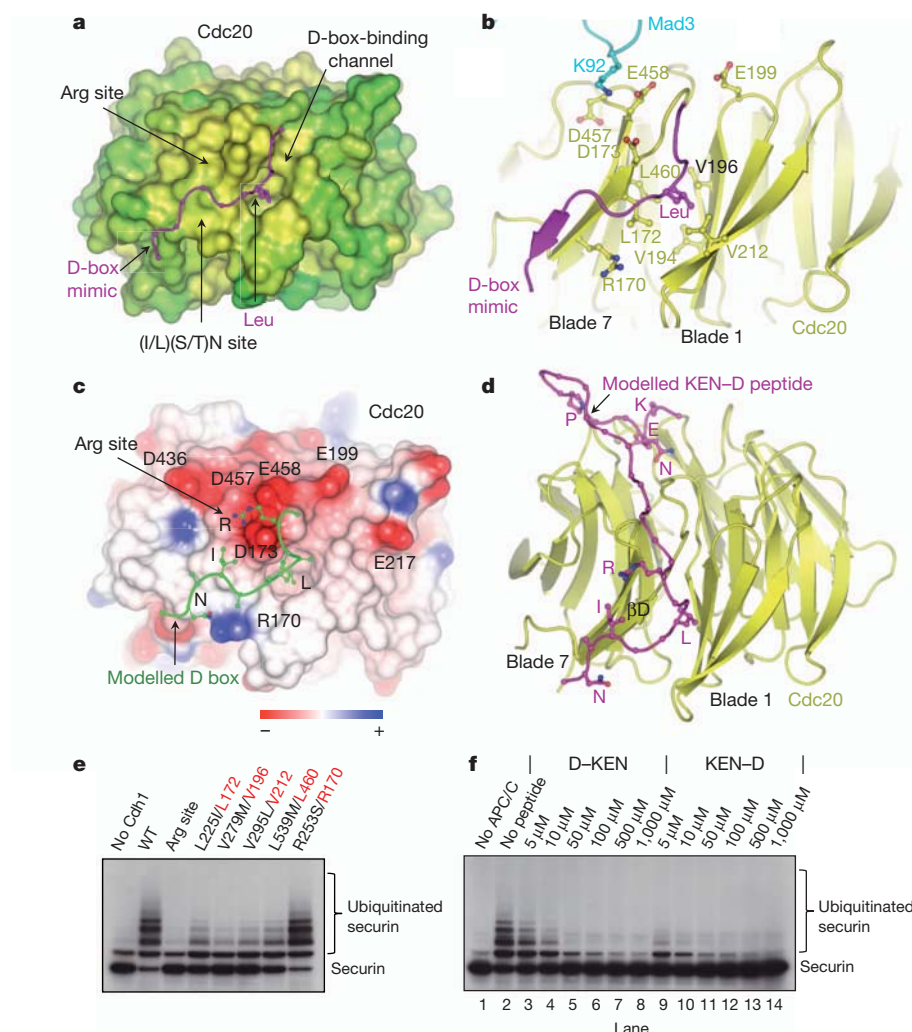


Figure 3 | The KEN box binds to a conserved surface at the centre of the top side of the WD40 domain. **a**, Cdc20 shown with surface representation coloured according to the sequence conservation of Cdc20 and Cdh1 (yellow (invariant) to green (less conserved)) (Supplementary Fig. 5). Mad3 diagram with KEN box in red. **b**, Details of KEN-box interactions with the Cdc20 WD40 domain. **c**, Mutation of the KEN-box-binding site abolishes *S. cerevisiae* APC/C^{Cdh1}-mediated ubiquitination of securin as indicated by an *in vitro* transcription or translation (IVT)-based ubiquitination assay. KEN-box mutants: N405A, N407A, Q473A, R517L (*S. pombe* Cdc20: N324, N326, Q392, R438).

Figure 4 | The D box binds in an extended conformation to a conserved inter-blade channel on the Cdc20 WD40 domain. **a**, Cdc20 shown with surface representation coloured according to sequence conservation of Cdc20 and Cdh1 (yellow (invariant) to green (less conserved)). (Supplementary Fig. 5). D-box mimic (C terminus of Mad3) is shown in purple. **b**, D-box mimic bound through a peptide-in-channel mechanism to Cdc20. Lys 92 of Mad3 contacts the negatively charged D-box Arg-binding site of Cdc20, sterically hindering the D-box receptor of Cdc20. **c**, Negative electrostatic potential of Cdc20 in the vicinity of the Arg-binding site of modelled D box. **d**, Model of a KEN-D peptide bound to KEN and D-box sites of Cdc20. **e**, Mutation of the D-box-binding site abolishes *S. cerevisiae* APC/C^{Cdh1}-mediated ubiquitination of securin as indicated by an IVT-based ubiquitination assay. D-box Arg-binding-site mutants: D256S, D536S, E537S (*S. pombe* Cdc20: D173, D457, E458). Mutated *S. cerevisiae* Cdh1 residues are indicated in black with equivalent residues of *S. pombe* in red. At the D-box Leu-binding site, mutation of *S. pombe* Cdc20 Val 196 had the most notable effect, with little loss of activity for Leu460Met. Mutation of Arg 170, located distal to the Arg and Leu sites (Fig. 4c), did not affect APC/C^{Cdh1} activity. **f**, Ubiquitination competition assay. The KEN-D box peptide inhibits *S. cerevisiae* APC/C^{Cdh1}-mediated securin ubiquitination fivefold more efficiently than does the D-box-KEN-box peptide (compare lanes 5 and 10). Blocking the Leu-binding pocket by small molecules ligands may be a potential mechanism to inhibit APC/C-co-activator activity.



equivalents of the Cdc20 residues Asn 326, Thr 368, Gln 392 and Arg 438 in *Saccharomyces cerevisiae* Cdh1 abolished APC/C^{Cdh1}-catalysed ubiquitination of the KEN-box and D-box-dependent APC/C substrates securin and Clb2 (Fig. 3c and Supplementary Fig. 6). Co-migration of APC/C and Cdh1 on a native gel confirmed that the inactivation was not due to misfolding of the mutant co-activator (Supplementary Fig. 7). Our findings that the KEN box of Mad3 is embedded within a segment of ordered structure, possibly stabilized on the Mad3 HLH motif through contacts to Mad2 (Fig. 1), suggest that degons do not need to be disordered to allow substrate recognition by ubiquitin ligases.

The WD40 domain of co-activators interacts with the D-box degon (RXXLXX(I/L)(S/T)N)^{25,39}, acting as a co-receptor with the APC/C subunit Apc10 (refs 40, 41). However, the molecular details of co-activator-D box interactions have not been established. In the crystal structure of the MCC, an unexpected crystal packing contact, in which the C terminus of a neighbouring Mad3 subunit partially mimics a D box and contacts Cdc20, provides detailed molecular insights of co-activator-D box recognition. The D box mimic of Mad3 was found in an extended conformation engaged within a conserved channel located between blades 1 and 7 on the rim of the β -propeller toroid (Fig. 4b and Supplementary Fig. 8). The interaction is dominated by the burial of the aliphatic side chain of Leu 215 of Mad3, mimicking the Leu of the D-box RXXL motif, within a deep pocket perfectly matching a Leu side chain, and created by non-polar residues invariant in Cdc20 and Cdh1 (Fig. 4a and Supplementary Fig. 5). Other consensus residues of the D box—the essential and

invariant Arg and less-well-conserved (I/L)(S/T)N motif⁴²—are not represented in the Mad3 D-box mimic. However, a candidate recognition site for the Arg side chain of the D box can be assigned to a cluster of negatively charged residues (Asp 173, Asp 457 and Glu 458) located on two adjacent β -strands of blade 7 on the top side of the D-box-binding channel. These residues are suitably positioned to interact with an Arg residue that is N-terminal to the anchored Leu residue, and are conserved in Cdh1 (Fig. 4b, c and Supplementary Fig. 5).

To test whether the putative D-box-binding site that was identified on Cdc20 is responsible for recognizing the D box of APC/C substrates, we generated mutants to abolish individually the equivalent Arg- and Leu-binding sites in Cdh1. Disrupting the Leu-binding pocket by replacing Val 196 with a bulky methionine, and substituting Ser for the Cdh1 equivalents of Asp 173, Asp 457 and Glu 458 at the putative Arg-binding site, eliminated the ability of Cdh1 to stimulate APC/C activity (Fig. 4e and Supplementary Fig. 6) but had no effect on co-activator binding to the APC/C (Supplementary Fig. 7). These results are consistent with the inter-blade channel of Cdc20 functioning as the D-box-binding site. The region of this channel that probably contacts the D-box (I/L)(S/T)N motif is less well conserved between Cdc20 and Cdh1 (Fig. 4a), possibly explaining the differential affinity of the two co-activators for D-box degons.

Substrate and inhibitor recognition by co-activators

To characterize further the APC/C degon-binding sites on co-activators, we designed peptides that incorporate both KEN-box and D-box motifs and tested their ability to inhibit APC/C^{Cdh1}.

Examination of the Cdc20 structure suggested that a peptide with KEN and D boxes linked by 17 residues (KEN–D peptide) would enable cooperative binding of both degrons to their respective binding sites, thereby conferring higher-affinity binding than peptides with either degron alone (Fig. 4d). In contrast, a peptide with the reverse orientation of KEN and D boxes (D–KEN peptide), also connected by a 17-residue linker, would only permit a single degron to bind. When we compared the potential of these peptides to competitively inhibit securin ubiquitination by APC/C^{Cdh1}, we found that the KEN–D peptide was five times more potent as an inhibitor than the D–KEN peptide (Fig. 4f). Because D–KEN peptide inhibited APC/C^{Cdh1} with the same efficiency as a D-box peptide (Supplementary Fig. 9a), these results indicate that cooperative degron binding to Cdh1 was conferred by the specific spatial arrangement of the KEN and D boxes of the KEN–D peptide, in agreement with our assignment of the KEN- and D-box-binding sites on co-activators.

The spatial arrangement of KEN and D boxes in the KEN–D inhibitory peptide is markedly similar to the organization of KEN and D boxes in two APC/C inhibitors: Acm1 and Mes1 (refs 43–45) (Supplementary Fig. 10). Both proteins inhibit the APC/C co-activators through a pseudosubstrate-based mechanism that is dependent on their D and KEN boxes, and it seems reasonable to assume that the spacing of 18 and 24 residues between the KEN and D box degrons of Acm1 and Mes1, respectively, optimizes inhibitor-co-activator affinity. Here we show one mechanism by which a KEN- and D-box-containing protein would bind co-activators. However, because of the diverse configuration and relative separation of KEN- and D-box motifs in APC/C substrates, several modes of co-activator–substrate recognition probably exist. Our assignment of the D-box recognition site to an inter-blade channel of the co-activator β -propeller was recently confirmed by a Cdh1–Acm1 crystal structure (W.C.H.C., D.B. and J. He, unpublished observations).

KEN- and D-box motifs adopt different conformations when bound to Cdc20; the KEN box is an underwound helix (Fig. 3b), whereas the D box assumes an extended structure (Fig. 4b). This important distinction means that the three residues of the KEN motif are presented to the same surface of Cdc20, whereas in the D box alternative amino acid side chains are oriented in opposite directions. Thus, with the Leu side chain of the D box anchored by the co-activator, conserved Arg, Ile/Leu and Asn side chains at positions 1, 7 and 9 of the D box⁴², respectively, would be accessible to generate a composite D-box-co-activator recognition surface for the D-box co-receptor Apc10 (refs 40, 41). This is consistent with the identification of the D-box co-receptor at the interface of co-activator and Apc10 in an APC/C^{Cdh1-D box} ternary complex⁴¹.

Implications for MCC-mediated inhibition of the APC/C

The processes underlying MCC-mediated inhibition of APC/C are incompletely defined, and probably involve several mechanisms. Mad3, which is dependent on its N-terminal KEN box, blocks Cdc20-mediated substrate recognition¹⁵, consistent with our structure showing that the Mad3 KEN box binds to the KEN-box recognition site of Cdc20. Notably, the same KEN box also promotes Mad3-dependent APC/C-mediated degradation of Cdc20 (refs 18, 27, 29), which suggests that Mad3 has a role in positioning Cdc20 for ubiquitination by the APC/C's catalytic centre. To understand this function of Mad3, we docked our MCC coordinates into the electron-microscope-derived molecular envelope of the APC/C–MCC complex¹⁴ (Fig. 5). Interpretation of this structure was based on our previous subunit assignment and pseudo-atomic model of budding yeast APC/C^{Cdh1-D box} (ref. 46). The MCC crystal structure corresponds closely to the assigned MCC density of APC/C^{MCC}, with Cdc20, Mad2 and Mad3 clearly recognizable. Mad2 contacts the TPR subunits Cdc23 and Apc5, whereas Mad3 interacts with Apc1. There is insufficient unassigned density to account for the C-terminal kinase domain of BubR1 and Bub3, suggesting their structural disorder. Notably, the

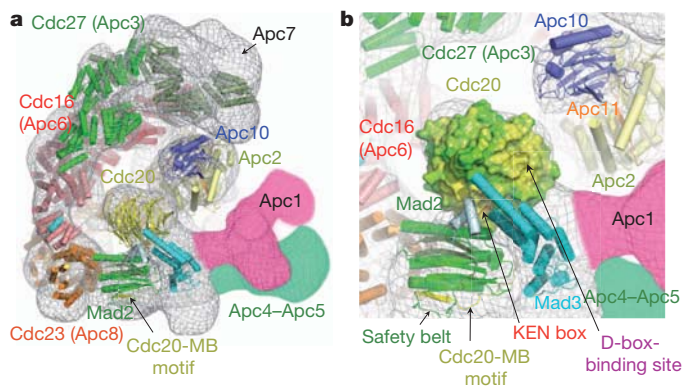


Figure 5 | Pseudo-atomic structure of human APC/C^{MCC}. **a**, Docking of MCC coordinates into a human APC/C^{MCC} electron microscope map¹⁴. APC/C subunit coordinates are based on the pseudo-atomic structure of *S. cerevisiae* APC/C⁴⁶. In the context of the APC/C^{MCC}, Cdc20 is in a downwards position relative to APC/C^{Cdh1} (ref. 41). **b**, Details of MCC interactions with APC/C showing the position of the D-box co-receptor on Cdc20 relative to Apc10 and the catalytic centre of Apc2 and Apc11. The downwards position of Cdc20 prevents interaction with Apc10. Surface representation of Cdc20 is colour-coded from yellow (invariant) to green (less conserved).

KEN-box-binding site of Cdc20 is blocked by Mad3, whereas the D-box site is directed towards, but not in contact with, its co-receptor Apc10 (Fig. 5b). Compared with APC/C^{Cdh1-D box}, in the APC/C^{MCC} complex, Cdc20 is displaced downwards towards Apc5. This position may facilitate Cdc20 ubiquitination. Furthermore, the lower position of co-activator prevents its D-box-binding site from generating a bipartite D-box co-receptor with Apc10 (ref. 41). This, together with a partial steric blockade of the arginine site of the D-box co-receptor of Cdc20 by Lys 92 of the Mad3 TPR domain (Fig. 4b), explains the inability of APC/C^{MCC} to recognize and ubiquitinate D-box-dependent substrates¹⁴. In APC/C^{MCC}, a region of the Cdc20 D-box-binding site for the Leu anchor and C-terminal residues, is accessible (Fig. 5b), suggesting that a non-consensus D-box sequence (that is, lacking an Arg residue and not dependent on the co-receptor Apc10) could engage this site.

In addition to the interactions of Mad2 and Mad3 with APC/C subunits, the position and activity of Cdc20 in the context of APC/C^{MCC} might also be influenced by the sequestration of its N-terminal APC/C recognition motifs by Mad2. The Mad2-binding motif of Cdc20 also mediates APC/C interactions⁴⁷, which would be blocked by the Mad2 safety belt (Fig. 5b). Furthermore, constraining the Cdc20 N terminus by Mad2 might also prevent the neighbouring C box from accessing its APC/C-binding site, which is necessary for Cdc20-dependent stimulation of APC/C catalytic activity⁴⁸.

The SAC is antagonized by p31^{comet} (refs 30, 49), and the structure of the C-Mad2–p31^{comet} complex showed that p31^{comet} binds at the Mad2 dimerization interface to inhibit the conformational activation of O-Mad2 (ref. 31). However, it is clear from our MCC structure that because Mad3 and p31^{comet} bind to a common Mad2 interface, p31^{comet} would compete for Mad3 interactions with Mad2 (Fig. 2c), explaining how p31^{comet} both antagonizes the assembly of the MCC³¹ and promotes its disassembly^{32,50}.

Here we show how the molecular basis for APC/C inhibition by its regulators is coupled to degron recognition by co-activators. APC/C activity is modulated by sterically blocking substrate recognition and through conformational changes that disrupt the substrate-binding site, reminiscent of modes of protein kinase regulation. This contrasts with the SCF complex whose activity is regulated at the level of substrate recognition through degron phosphorylation.

METHODS SUMMARY

Expression, purification and crystallization of *S. pombe* MCC. The *S. pombe* MCC (involving Cdc20, Mad2 and Mad3) was generated by co-expressing

Cdc20-Slp1 (residues 87–488), Mad2 (with Leu12Ala and Arg133Ala mutations) and Mad3 (residues 1–223) in the baculovirus and insect cell systems. Crystallization, crystal structure determination, mutagenesis, enzyme assays and other procedures were performed as described in the Methods. Diffraction data were collected at beamline I02 at the Diamond Light Source.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 September 2011; accepted 25 January 2012.

Published online 21 March 2012.

- Musacchio, A. & Salmon, E. D. The spindle-assembly checkpoint in space and time. *Nature Rev. Mol. Cell Biol.* **8**, 379–393 (2007).
- Kim, S. & Yu, H. Mutual regulation between the spindle checkpoint and APC/C. *Semin. Cell Dev. Biol.*, (2011).
- Hoyt, M. A., Totis, L. & Roberts, B. T. S. *cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. *Cell* **66**, 507–517 (1991).
- Li, R. & Murray, A. W. Feedback control of mitosis in budding yeast. *Cell* **66**, 519–531 (1991).
- Li, Y., Gorbea, C., Mahaffey, D., Rechsteiner, M. & Benezra, R. MAD2 associates with the cyclosome/anaphase-promoting complex and inhibits its activity. *Proc. Natl Acad. Sci. USA* **94**, 12431–12436 (1997).
- Fang, G., Yu, H. & Kirschner, M. W. The checkpoint protein MAD2 and the mitotic regulator CDC20 form a ternary complex with the anaphase-promoting complex to control anaphase initiation. *Genes Dev.* **12**, 1871–1883 (1998).
- Hwang, L. H. et al. Budding yeast Cdc20: a target of the spindle checkpoint. *Science* **279**, 1041–1044 (1998).
- Kim, S. H., Lin, D. P., Matsumoto, S., Kitazono, A. & Matsumoto, T. Fission yeast Slp1: an effector of the Mad2-dependent spindle checkpoint. *Science* **279**, 1045–1047 (1998).
- Kallio, M., Weinstein, J., Daum, J. R., Burke, D. J. & Gorbsky, G. J. Mammalian p55CDC mediates association of the spindle checkpoint protein Mad2 with the cyclosome/anaphase-promoting complex, and is involved in regulating anaphase onset and late mitotic events. *J. Cell Biol.* **141**, 1393–1406 (1998).
- Hardwick, K. G., Johnston, R. C., Smith, D. L. & Murray, A. W. MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p. *J. Cell Biol.* **148**, 871–882 (2000).
- Tang, Z., Bharadwaj, R., Li, B. & Yu, H. Mad2-Independent inhibition of APC^{Cdc20} by the mitotic checkpoint protein BubR1. *Dev. Cell* **1**, 227–237 (2001).
- Fang, G. Checkpoint protein BubR1 acts synergistically with Mad2 to inhibit anaphase-promoting complex. *Mol. Biol. Cell* **13**, 755–766 (2002).
- Sudakin, V., Chan, G. K. & Yen, T. J. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *J. Cell Biol.* **154**, 925–936 (2001).
- Herzog, F. et al. Structure of the anaphase-promoting complex/cyclosome interacting with a mitotic checkpoint complex. *Science* **323**, 1477–1481 (2009).
- Burton, J. L. & Solomon, M. J. Mad3p, a pseudosubstrate inhibitor of APC^{Cdc20} in the spindle assembly checkpoint. *Genes Dev.* **21**, 655–667 (2007).
- Davenport, J., Harris, L. D. & Goorha, R. Spindle checkpoint function requires Mad2-dependent Cdc20 binding to the Mad3 homology domain of BubR1. *Exp. Cell Res.* **312**, 1831–1842 (2006).
- Kulukian, A., Han, J. S. & Cleveland, D. W. Unattached kinetochores catalyze production of an anaphase inhibitor that requires a Mad2 template to prime Cdc20 for BubR1 binding. *Dev. Cell* **16**, 105–117 (2009).
- Nilsson, J., Yekezare, M., Minshull, J. & Pines, J. The APC/C maintains the spindle assembly checkpoint by targeting Cdc20 for destruction. *Nature Cell Biol.* **10**, 1411–1420 (2008).
- Luo, X. & Yu, H. Protein metamorphosis: the two-state behavior of Mad2. *Structure* **16**, 1616–1625 (2008).
- Luo, X. et al. Structure of the Mad2 spindle assembly checkpoint protein and its interaction with Cdc20. *Nature Struct. Biol.* **7**, 224–229 (2000).
- Luo, X., Tang, Z., Rizo, J. & Yu, H. The Mad2 spindle checkpoint protein undergoes similar major conformational changes upon binding to either Mad1 or Cdc20. *Mol. Cell* **9**, 59–71 (2002).
- Sironi, L. et al. Crystal structure of the tetrameric Mad1–Mad2 core complex: implications of a ‘safety belt’ binding mechanism for the spindle checkpoint. *EMBO J.* **21**, 2496–2506 (2002).
- De Antoni, A. et al. The Mad1/Mad2 complex as a template for Mad2 activation in the spindle assembly checkpoint. *Curr. Biol.* **15**, 214–225 (2005).
- Barford, D. Structure, function and mechanism of the anaphase promoting complex (APC/C). *Q. Rev. Biophys.* **44**, 153–190 (2011).
- Glotzer, M., Murray, A. W. & Kirschner, M. W. Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132–138 (1991).
- Pfleger, C. M. & Kirschner, M. W. The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1. *Genes Dev.* **14**, 655–665 (2000).
- King, E. M., van der Sar, S. J. & Hardwick, K. G. Mad3 KEN boxes mediate both Cdc20 and Mad3 turnover, and are critical for the spindle checkpoint. *PLoS ONE* **2**, e342 (2007).
- Sczaniecka, M. et al. The spindle checkpoint functions of Mad3 and Mad2 depend on a Mad3 KEN box-mediated interaction with Cdc20-anaphase-promoting complex (APC/C). *J. Biol. Chem.* **283**, 23039–23047 (2008).
- Pan, J. & Chen, R. H. Spindle checkpoint regulates Cdc20p stability in *Saccharomyces cerevisiae*. *Genes Dev.* **18**, 1439–1451 (2004).
- Habu, T., Kim, S. H., Weinstein, J. & Matsumoto, T. Identification of a MAD2-binding protein, CMT2, and its role in mitosis. *EMBO J.* **21**, 6419–6428 (2002).
- Yang, M. et al. p31^{comet} blocks Mad2 activation through structural mimicry. *Cell* **131**, 744–755 (2007).
- Reddy, S. K., Rape, M., Margansky, W. A. & Kirschner, M. W. Ubiquitination by the anaphase-promoting complex drives spindle checkpoint inactivation. *Nature* **446**, 921–925 (2007).
- D’Arcy, S., Davies, O. R., Blundell, T. L. & Bolanos-Garcia, V. M. Defining the molecular basis of BubR1 kinetochore interactions and APC/C-CDC20 inhibition. *J. Biol. Chem.* **285**, 14764–14776 (2010).
- Malureanu, L. A. et al. BubR1 N terminus acts as a soluble inhibitor of cyclin B degradation by APC/C(Cdc20) in interphase. *Dev. Cell* **16**, 118–131 (2009).
- Schwab, M., Neutzner, M., Mockner, D. & Seufert, W. Yeast Hct1 recognizes the mitotic cyclin Clb2 and other substrates of the ubiquitin ligase APC. *EMBO J.* **20**, 5165–5175 (2001).
- Vodermaier, H. C., Gieffers, C., Maurer-Stroh, S., Eisenhaber, F. & Peters, J. M. TPR subunits of the anaphase-promoting complex mediate binding to the activator protein CDH1. *Curr. Biol.* **13**, 1459–1468 (2003).
- Mapelli, M., Massimiliano, L., Santaguida, S. & Musacchio, A. The Mad2 conformational dimer: structure and implications for the spindle assembly checkpoint. *Cell* **131**, 730–743 (2007).
- Tipton, A. R. et al. BUBR1 and closed MAD2 (C-MAD2) interact directly to assemble a functional mitotic checkpoint complex. *J. Biol. Chem.* **286**, 21173–21179 (2011).
- Kraft, C., Vodermaier, H. C., Maurer-Stroh, S., Eisenhaber, F. & Peters, J. M. The WD40 propeller domain of Cdh1 functions as a destruction box receptor for APC/C substrates. *Mol. Cell* **18**, 543–553 (2005).
- Carroll, C. W., Enquist-Newman, M. & Morgan, D. O. The APC subunit Doc1 promotes recognition of the substrate destruction box. *Curr. Biol.* **15**, 11–18 (2005).
- da Fonseca, P. C. et al. Structures of APC/C(Cdh1) with substrates identify Cdh1 and Apc10 as the D-box co-receptor. *Nature* **470**, 274–278 (2011).
- King, R. W., Glotzer, M. & Kirschner, M. W. Mutagenic analysis of the destruction signal of mitotic cyclins and structural characterization of ubiquitinated intermediates. *Mol. Biol. Cell* **7**, 1343–1357 (1996).
- Choi, E., Dial, J. M., Jeong, D. E. & Hall, M. C. Unique D box and KEN box sequences limit ubiquitination of Acm1 and promote pseudosubstrate inhibition of the anaphase-promoting complex. *J. Biol. Chem.* **283**, 23701–23710 (2008).
- Ostapenko, D., Burton, J. L., Wang, R. & Solomon, M. J. Pseudosubstrate inhibition of the anaphase-promoting complex by Acm1: regulation by proteolysis and Cdc28 phosphorylation. *Mol. Cell Biol.* **28**, 4653–4664 (2008).
- Kimata, Y. et al. A mutual inhibition between APC/C and its substrate Mes1 required for meiotic progression in fission yeast. *Dev. Cell* **14**, 446–454 (2008).
- Schreiber, A. et al. Structural basis for the subunit assembly of the anaphase-promoting complex. *Nature* **470**, 227–232 (2011).
- Zhang, Y. & Lees, E. Identification of an overlapping binding domain on Cdc20 for Mad2 and anaphase-promoting complex: model for spindle checkpoint regulation. *Mol. Cell Biol.* **21**, 5190–5199 (2001).
- Kimata, Y., Baxter, J. E., Fry, A. M. & Yamano, H. A role for the Fizzy/Cdc20 family of proteins in activation of the APC/C distinct from substrate recruitment. *Mol. Cell* **32**, 576–583 (2008).
- Xia, G. et al. Conformation-specific binding of p31(comet) antagonizes the function of Mad2 in the spindle checkpoint. *EMBO J.* **23**, 3133–3143 (2004).
- Westhorpe, F. G., Tighe, A., Lara-Gonzalez, P. & Taylor, S. S. p31comet-mediated extraction of Mad2 from the MCC promotes efficient mitotic exit. *J. Cell Sci.* **124**, 3905–3916 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by a Cancer Research UK grant to D.B. and an ICR studentship to W.C.H.C. We thank J. Yang and J. He for advice, and staff at the Diamond Light Source beamline I02 for help with data collection.

Author Contributions W.C.H.C. cloned, purified and crystallized the *S. pombe* MCC, and performed the biochemistry and mutagenesis experiments. W.C.H.C., K.K. and D.B. collected the X-ray diffraction data. K.K. determined the complex structure and modelled the D-box and KEN-D peptides. Z.Z. advised on cloning strategies of the MCC. E.H.K. provided purified endogenous APC/C. W.C.H.C., K.H.K. and D.B. docked crystal coordinates into the human APC/C^{MCC} electron-microscope map. W.C.H.C. and D.B. wrote the manuscript and the others authors provided editorial input.

Author Information The MCC coordinates have been deposited in the Protein Data Bank under accession number 4aez. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.B. (david.barford@icr.ac.uk).

METHODS

Cloning, expression and purification of the MCC complex.

Schizosaccharomyces pombe CDC20-SLP1 (residues 87–488), MAD2 (residues 1–203) and MAD3 (residues 1–223) genes were amplified by polymerase chain reaction (PCR) using an *S. pombe* complementary DNA library pTN-TH7 (a gift from T. Nakamura) as a template and cloned into a modified pFBDM vector⁵¹ (Z.Z. & D.B., unpublished observations). A double Strep-tag II (ds) together with a tobacco etch virus (TEV) cleavage site were introduced into the N terminus of Cdc20-Slp1 and mutations of Leu12Ala and Arg133Ala were introduced into Mad2 (ref. 31). A previous study showed that the Arg133Ala mutant was fully capable of inducing a mitotic arrest⁵². The resultant protein expression cassettes were recombined with the DH10MultiBac cells to create a bacmid by transposition⁵¹. The ds-Cdc20–Mad2–Mad3 complex was expressed using the baculovirus and insect cell (High 5 cells) systems, and purified by a combination of Strep-Tactin (Qiagen), anion exchange chromatography Resource Q and Superdex 200 size-exclusion chromatography (GE Healthcare). The identities of the purified complex subunits were confirmed by antibody labelling and N-terminal sequencing. Mutagenesis was performed to generate the Mad2–Mad3 interface mutant, Cdc20–Mad2(A137Y)–Mad3(M13R).

Crystallization, data collection and structure determination. Crystals were obtained in a hanging-drop fashion by pre-incubating 1:1 v/v of 4.5 mg ml^{−1} of protein with the crystallization solution, 100 mM Tris-HCl pH 8.8, 21% PEG 3350, 30% ethylene glycol, 5 mM dithiothreitol (DTT), and 5 mM EDTA at 20 °C for 24 h, followed by streak seeding. Crystals grew to full size after 2 weeks and were mounted in 0.2–0.3 mm cryoloops and frozen in liquid nitrogen. Native crystals diffracted to a minimum Bragg spacing (d_{\min}) of about 2.1 Å. The diffraction data set was collected at the I02 beamline of Diamond Light Source from a single crystal, processed with XDS and scaled to 2.3 Å with SCALA⁵³. Phase information was obtained by molecular replacement with AMoRe⁵⁴. Monomeric coordinates from the crystal structures of human MAD2(L13A) (PDB 2VFX)⁵⁵, human BUBR1 (PDB 2WV1)³³ and human WDR5 (PDB 3EMH)⁵⁶ were used as search models. All three trimeric complexes (MCC complexes) that are present in the asymmetric unit were assembled with repeated runs of AMoRe. Iterative model building and refinements were carried out with COOT⁵⁷ and PHENIX⁵⁸, respectively. Simulated annealing composite omit maps were systematically calculated at several stages of model building and refinement, and examined to minimize the effects of model bias. TLS parameters that were generated from the TLS motion determination server⁵⁹ (<http://skuld.bmsc.washington.edu/~tlsmd/>) were used throughout the refinement. Water molecules were added towards the end of the refinement. The structure was validated with MolProbity⁶⁰. Data collection and refinement statistics are given in Supplementary Table 1.

APC/C ubiquitination assays with wild-type and mutant Cdh1. *Saccharomyces cerevisiae* Cdh1 mutants were generated using PCR-based mutagenesis and cloned into a linearized pRSET vector. A functional T7 promoter and the Cdh1 sequences were confirmed by DNA sequencing. APC/C ubiquitination assays were adopted and modified from ref. 61. ³⁵S-labelled Clb2p and securin (Pds1p) and unlabelled Cdh1 mutants were prepared using TNT T7 Quick-coupled *in vitro* transcription (or translation) (IVT; Promega). Each ubiquitination reaction contains approximately 10 ng of *S. cerevisiae* APC/C, 1 µl of ³⁵S-labelled substrate and 2 µl of Cdh1 in a 10-µl reaction volume with 40 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 0.6 mM DTT, 2.7 mM ATP, 6.6 µg of methyl ubiquitin, 500 ng of Ubc4, 200 ng of ubiquitin aldehyde (Enzo Life Science), 2 µM

LLnL (*N*-acetyl-Leu-Leu-Norleu-aldehyde) (Sigma). Reactions were incubated at room temperature for 15 min and were analysed using 8% SDS-PAGE. Gels were fixed and stained with Coomassie blue, then dried and exposed to BioMax MR Film (Kodak).

Native gel electrophoresis. Correct folding of *S. cerevisiae* Cdh1 mutants was assessed by their co-migration with the APC/C in native gel electrophoresis. 50 ng of apoAPC/C was mixed with 2 µl of ³⁵S-labelled IVT-produced Cdh1 and 0.7 µl of 100 mM CaCl₂ in a volume of 14 µl with 10 mM Tris pH 8.0, 150 mM NaCl, 3 mM DTT, 1 mM magnesium acetate and 2 mM EGTA. Samples were incubated at room temperature for 15 min before adding 1 µl of native gel loading buffer (125 mM Tris pH 8.0, 84% (v/v) glycerol) to each reaction. The entire reaction was loaded onto a 5.25% non-denaturing polyacrylamide gel run at 4 °C, 110 V for 2 h. Gels were fixed and stained with Coomassie blue, then dried and exposed to film.

KEN-D peptide competition assays. The KEN-D (Ac-NKENEGPASGASGASGASGAQRAALSDITNS-NH₂), D-KEN (Ac-QRAALSDITNSGASGASGASGASGASGASGAQRAALSDITNS-NH₂), KEN-D^{mut} (Ac-NKENEGPASGASGASGASGAQSAASDITSS-NH₂) and KEN^{mut}-D (Ac-NSASEGPASGASGASGASGAQRAALSDITNS-NH₂) peptides were designed with a 17-residue linker between the KEN-box sequence (NKENEGPA) and the D-box sequence (QRAALSDITNS), or between their respective mutant sequences (NSASEGPA and QSAASDITSS). The peptides were synthesized by Cambridge Peptides. APC/C ubiquitination reactions were performed in the same way as described above using the co-purified APC/C^{Cdh1} complex⁴¹. Peptides were added into the reaction to a final concentration ranging from 5 µM to 1000 µM. Reactions were incubated at room temperature for 15 min and were analysed using 8% SDS-PAGE. Gels were fixed and stained with Coomassie blue, then dried and exposed to film.

Fitting structure coordinates into the human APC/C^{MCC} EM map. The structure coordinates of *S. pombe* Cdc16–Cdc26 complex (PDB 2XPI), Cdc27 (PDB 3KAE), Apc10 (PDB 1GQP), modelled Apc2 and MCC were fitted into the human APC/C^{MCC} map (EMD-1591) based on the *S. cerevisiae* APC/C assignment⁴⁶, using UCSF Chimera.

- Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nature Biotechnol.* **22**, 1583–1587 (2004).
- Sironi, L. *et al.* Mad2 binding to Mad1 and Cdc20, rather than oligomerization, is required for the spindle checkpoint. *EMBO J.* **20**, 6371–6382 (2001).
- Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
- Navaza, J. Implementation of molecular replacement in AMoRe. *Acta Crystallogr. D* **57**, 1367–1372 (2001).
- Yang, M. *et al.* Insights into mad2 regulation in the spindle checkpoint revealed by the crystal structure of the symmetric mad2 dimer. *PLoS Biol.* **6**, e50 (2008).
- Song, J. J. & Kingston, R. E. WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J. Biol. Chem.* **283**, 35258–35264 (2008).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
- Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* **62**, 439–450 (2006).
- Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
- Passmore, L. A., Barford, D. & Harper, J. W. Purification and assay of the budding yeast anaphase-promoting complex. *Methods Enzymol.* **398**, 195–219 (2005).

Structural basis of ultraviolet-B perception by UVR8

Di Wu^{1*}, Qi Hu^{1*}, Zhen Yan^{1*}, Wen Chen², Chuangye Yan³, Xi Huang^{2,4}, Jing Zhang¹, Panyu Yang^{2,4}, Haiteng Deng¹, Jiawei Wang³, XingWang Deng^{2,4} & Yigong Shi¹

The *Arabidopsis thaliana* protein UVR8 is a photoreceptor for ultraviolet-B. Upon ultraviolet-B irradiation, UVR8 undergoes an immediate switch from homodimer to monomer, which triggers a signalling pathway for ultraviolet protection. The mechanism by which UVR8 senses ultraviolet-B remains largely unknown. Here we report the crystal structure of UVR8 at 1.8 Å resolution, revealing a symmetric homodimer of seven-bladed β -propeller that is devoid of any external cofactor as the chromophore. Arginine residues that stabilize the homodimeric interface, principally Arg 286 and Arg 338, make elaborate intramolecular cation- π interactions with surrounding tryptophan amino acids. Two of these tryptophans, Trp 285 and Trp 233, collectively serve as the ultraviolet-B chromophore. Our structural and biochemical analyses identify the molecular mechanism for UVR8-mediated ultraviolet-B perception, in which ultraviolet-B radiation results in destabilization of the intramolecular cation- π interactions, causing disruption of the critical intermolecular hydrogen bonds mediated by Arg 286 and Arg 338 and subsequent dissociation of the UVR8 homodimer.

Perception of light is important for all kingdoms of life¹. Light regulates the circadian clock in worms and social activity in fruitflies^{2,3}. In plants, light is a major source of energy and regulates all major developmental and physiological processes^{4,5}. A wide wavelength range of light is sensed by specific families of photoreceptors: phytochrome for red and far red⁶; phototropin and cryptochrome for ultraviolet-A and blue^{7–11}; and UVR8 for ultraviolet-B (wavelength range 280–315 nm)^{4,12}. Except for UVR8, all other photoreceptors contain specific external cofactors as chromophores: bilin for phytochrome; FAD and MTHF for cryptochrome; and FMN for phototropin^{4,13}. It remains unclear whether UVR8 contains any external cofactor for ultraviolet-B perception.

UVR8, originally identified as a regulatory protein for ultraviolet-B-triggered signal transduction¹⁴, was recently shown to be a receptor for ultraviolet-B¹². Ultraviolet-B perception was thought to induce dissociation of the UVR8 homodimer, allowing its subsequent interaction with COP1 and transcriptional activation of ultraviolet-B-responsive genes^{12,15–17} (Fig. 1a). A number of tryptophan residues, particularly Trp 285, were shown to have an important role in ultraviolet-B-triggered signalling¹². Despite these advances, it remains unknown how ultraviolet-B is sensed by UVR8 or how ultraviolet-B perception leads to dissociation of the UVR8 homodimers.

Biochemical characterization of UVR8

The full-length, wild-type UVR8 and two variants, W285F and W285A, were purified to homogeneity. As reported¹², both wild type and W285F existed mainly as homodimers on SDS-polyacrylamide gel electrophoresis (SDS-PAGE) in the absence of heating (Fig. 1b, lanes 1–2). However, only wild type, not W285F, was able to undergo ultraviolet-B-induced monomerization (lanes 7–8). By contrast, the variant W285A appeared only as a monomer both before and after ultraviolet-B irradiation (lanes 3 and 9). Heating at 96 °C in the presence of SDS reduced all UVR8 homodimers to a monomeric state (lanes 4–6). These results are in agreement with published observations¹².

The ionic detergent SDS is a protein denaturant. Homodimer formation of wild-type UVR8, however, is remarkably stable and resists treatment by up to 12% SDS in the absence of heating (Supplementary Fig. 1a). Because the SDS sample buffer contains 200 mM dithiothreitol (DTT), we speculated that two molecules of UVR8 are held together through a covalent linkage—such as a disulphide bond—that is susceptible to heating in the presence of DTT or ultraviolet-B irradiation. Quite unexpectedly, elevated ionic strength led to efficient conversion of the UVR8 homodimers to a monomeric state (Supplementary Fig. 1b). This observation strongly suggests that the forces that hold together two UVR8 molecules are ionic in nature.

Use of SDS-PAGE may not allow faithful evaluation of native protein conformation. To alleviate this potential problem, we used the more sensitive method of gel filtration to examine the oligomeric state of UVR8 under physiological pH and ionic strength (Fig. 1c). Before ultraviolet-B irradiation, wild-type UVR8 was eluted from gel filtration with an apparent molecular mass of approximately 150 kDa, larger than that calculated for a UVR8 homodimer (~100 kDa). This discrepancy is probably caused by the extended flexible sequences at the amino and carboxy termini of UVR8. After ultraviolet-B irradiation, the elution volume for wild-type UVR8 corresponded to a molecular mass of about 75 kDa (Fig. 1c). This observation confirms the reported, ultraviolet-B-induced dimer-to-monomer switch¹². As anticipated, the variant UVR8(W285F) appeared as a homodimer irrespective of ultraviolet-B radiation (Fig. 1c). However, unlike the SDS-PAGE result¹² (Fig. 1b), UVR8(W285A) existed mainly as a homodimer both with and without ultraviolet-B treatment (Fig. 1c). Ultraviolet-B irradiation seemed to have weakened formation of the W285A homodimer as judged by the presence of a small fraction of monomers (Fig. 1c).

The WD40 repeats of UVR8 are thought to be responsible for ultraviolet-B perception¹². Supporting this notion, the protease-resistant core domain of UVR8 (residues 12–381; Supplementary

¹Tsinghua-Peking Center for Life Sciences, Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China. ²College of Life Sciences, Peking University, Beijing 100871, China. ³State Key Laboratory of Bio-membrane and Membrane Biotechnology, Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China. ⁴Peking-Yale Joint Center of Plant Molecular Genetics and Agrobiotechnology, State Key Laboratory of Protein and Plant Gene Research, Peking University, Beijing 100871, China. *These authors contributed equally to this work.

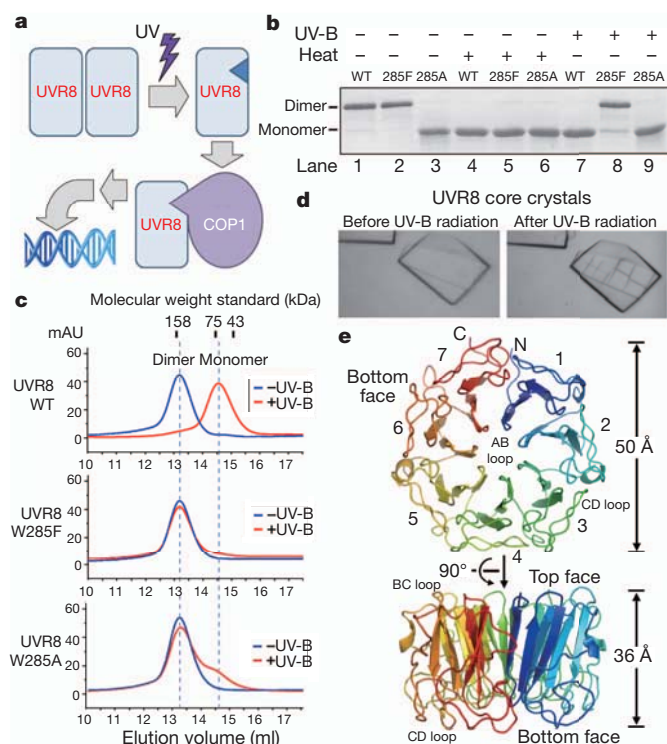


Figure 1 | Characterization and structure of the ultraviolet-sensing protein UVR8. **a**, A schematic diagram of ultraviolet-B-induced, UVR8-mediated signalling cascade. Upon receiving ultraviolet-B (UV-B) radiation, the UVR8 homodimer dissociates into monomers. The UVR8 monomer then associates with COP1, ultimately resulting in the activation of ultraviolet-B-responsive genes. **b**, The wild-type (WT) UVR8 undergoes a dimer-to-monomer switch in response to ultraviolet-B radiation as judged by SDS-PAGE. **c**, Solution behaviour of UVR8 in response to ultraviolet-B radiation. Shown here are gel filtration chromatograms. Only the wild-type UVR8, but not the variants W285F or W285A, switched from a homodimer to a monomer in response to ultraviolet-B radiation. **d**, Crystals of the UVR8 core domain (residues 12–381) were cracked by ultraviolet-B radiation. **e**, Overall structure of the UVR8 core domain. All structural figures were prepared with PyMOL⁴⁰.

Fig. 1c) retained the same ability as the full-length protein to undergo an ultraviolet-B-induced, dimer-to-monomer switch (Supplementary Fig. 1d). To elucidate the mechanism of ultraviolet-B perception by UVR8, we crystallized its core domain. Intriguingly, ultraviolet-B irradiation resulted in cracking of these crystals (Fig. 1d), indicating that the UVR8 core domain retained the ability to sense ultraviolet-B in the crystals. We also crystallized the core domain variants W285F and W285A. By contrast, the W285F and W285A crystals failed to crack even after prolonged ultraviolet-B irradiation (Supplementary Fig. 2), consistent with loss of ultraviolet-B responsiveness¹² (Fig. 1c).

Overall structure of UVR8

The crystal structure of the UVR8 core domain (residues 12–381), which represents 84% of the full-length UVR8 protein, was determined by selenium-based, single-wavelength anomalous dispersion (SAD). The atomic models of the UVR8 core domain and its variants W285F and W285A were refined at resolutions of 1.8, 2.0 and 1.8 Å, respectively (Supplementary Table 1 and Supplementary Fig. 3). The UVR8 core domain forms a seven-bladed β -propeller (Fig. 1e). In contrast to all previously determined structures of photoreceptor¹³, UVR8 does not contain any external cofactor as the chromophore. Unlike canonical WD-40 repeats¹⁸, each blade in UVR8 comprises only three β -strands, termed A, B and C¹⁹ (Supplementary Fig. 4a). An extended loop, designated loop CD, follows strand C in each blade. By convention¹⁹, loops AB and CD reside on the bottom face of the β -propeller whereas the BC loop is located on the top face (Fig. 1e).

A prominent sequence motif, GWRHT, is present in the AB loops of blades 5–7 (Supplementary Fig. 4a). The seven blades in UVR8 have a nearly identical main-chain conformation, which is similar to that of the cell cycle regulatory protein RCC1 (ref. 20; Supplementary Fig. 4b, c).

In the crystals, two molecules of the UVR8 core domain form a homodimer. Two surface patches of complementary charges are located on the bottom face of each core domain (Fig. 2a). The acidic surface patch comprises five negatively charged amino acids, Glu 43, Asp 44 and Glu 53 from blade 1 and Asp 96 and Asp 107 from blade 2. The basic patch contains four positively charged residues, Lys 252 from blade 5, Arg 286 from blade 6, and Arg 338 and Arg 354 from blade 7. The acidic and basic surface patches from one UVR8 core domain associate with the complementarily charged surface patches of another core domain to form a symmetric homodimer (Fig. 2b and Supplementary Fig. 5). This homodimeric interface, involving 2,566 Å² buried surface area, is mediated by 32 intermolecular hydrogen bonds.

Charged amino acids in the two prominent surface patches contribute a total of 20 intermolecular hydrogen bonds at the homodimeric interface (Fig. 2c). In blade 1 of one UVR8 molecule, Glu 43, Asp 44 and Glu 53 accept four charge-stabilized hydrogen bonds from Arg 338 and Arg 354 in blade 7 of the other UVR8 molecule. Arg 354 also donates a hydrogen bond to the carbonyl oxygen of residue 52 in blade 1. In blade 2, Asp 96 and Asp 107 mediate four intermolecular hydrogen bonds from Arg 286 of blade 6, whereas Ser 106 receives an intermolecular hydrogen bond from Lys 252 of blade 5. In addition to the two prominent surface patches, residues in blade 3 of one UVR8 molecule associate symmetrically with amino acids in blades 4 and 3 from the other UVR8 molecule, contributing 12 additional intermolecular hydrogen bonds (Fig. 2c, bottom right).

All 32 intermolecular hydrogen bonds involve side chains, of which 28 are mediated exclusively by side chains. Apart from two contacts between Gln 148 and Asn 149 (Fig. 2c, bottom right), all 30 other hydrogen bonds rely on charged amino acids and 24 are made between residues of opposite charges. Notably, Arg 286 of blade 4 contributes eight intermolecular hydrogen bonds (four from each molecule) at the homodimeric interface. This analysis explains the finding that homodimerization of UVR8 is disrupted by elevated ionic strength (Supplementary Fig. 1b).

Intramolecular cation– π interactions

The overall structures of the UVR8 variants W285F and W285A are nearly identical to that of the core domain, with a pair-wise root-mean-squared deviation of less than 0.5 Å for all aligned C α atoms (Fig. 3a, left). Analysis of the local structural features surrounding residue 285 reveals no significant conformational changes between the core domain and the variant W285F (Fig. 3a and Supplementary Fig. 6). The aromatic side chain of Phe 285 in W285F occupies the same position, with a similarly planar orientation, as Trp 285 in the core domain. By sharp contrast, major changes occur to three critical residues in the variant W285A (Fig. 3a and Supplementary Fig. 6). Compared to wild-type UVR8, the indole ring of Trp 337 swings approximately 4 Å to occupy the vacated space due to replacement of Trp 285 by Ala. Consequently, the indole ring of Trp 233 is rotated 180° around the C β –C γ axis, which in turn drives a 3.8-Å movement by the carboxylate group of Asp 129.

Cation– π interactions, known to stabilize protein structure²¹, make considerably more contribution to free energy terms than simple ionic interactions^{22,23}. Arginine and tryptophan are the most preferred amino acids for cation– π interactions²¹, with strong interactions for the distance range of 3.4–4.5 Å between the cation and the aromatic ring²². A detailed analysis of the UVR8 structure revealed an extensive network of intramolecular cation– π interactions surrounding Trp 285 and Arg 286 in blade 6 (Fig. 3b).

At the centre of the network, Arg 286 is surrounded by four aromatic amino acids (Fig. 3b). The guanidinium group of Arg 286 is positioned at approximately 3.8 and 4.2 Å away from the indole

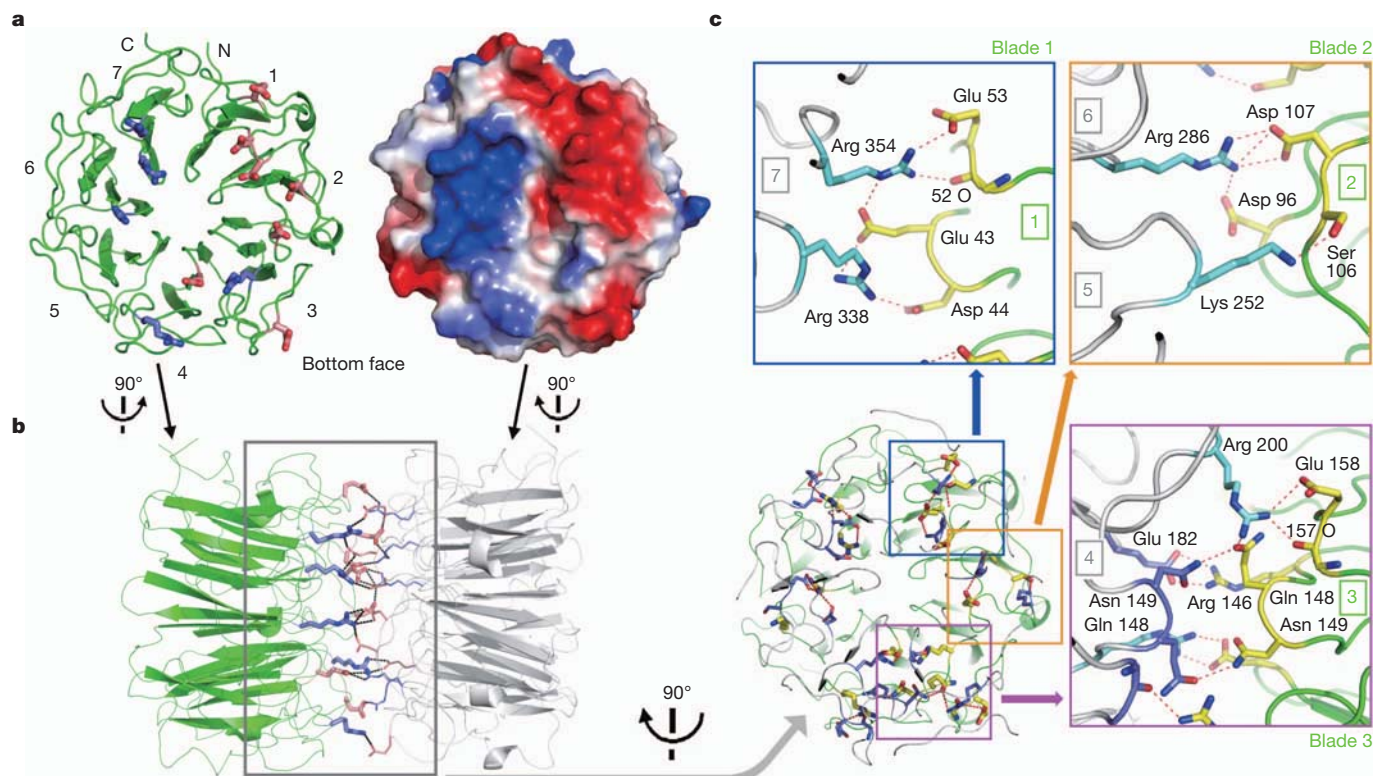


Figure 2 | Structural basis of UVR8 homodimer formation. **a**, The bottom face of the UVR8 propeller contains two surface patches of complementary charges. The left and right panels depict the cartoon and electrostatic surface representations of the UVR8 core domain, respectively. Five acidic amino acids (coloured red) from blades 1 and 2 constitute the negatively charged patch, whereas four basic residues (coloured blue) form the positively charged patch.

b, Formation of UVR8 homodimer is mediated by charge-stabilized hydrogen bonds, mainly from the two surface patches of complementary charges. The two UVR8 molecules are related to those in panel **a** by 90° rotations, as indicated in the figure. **c**, A close-up view on the intermolecular hydrogen bonds at the homodimeric interface. Hydrogen bonds are represented by red dashed lines.

rings of Trp 302 and Trp 285, respectively, allowing cation- π interactions of maximal strength²². Arg 286 also interacts with the phenyl ring of Tyr 253 and weakly associates with Trp 250. In addition, the indole ring of Trp 285 associates with the guanidinium group of Arg 338 through strong cation- π interactions, whereas the indole ring of Trp 285 binds to Trp 233 and Trp 337 through π - π stacking interactions²⁴ (Fig. 3b). The essence of π - π stacking is the cation- π interaction, as the edge of the tryptophan indole ring (around the amine group) carries net positive charges²³. Furthermore, Trp 233 appears to have an important role in this network by making π - π stacking interactions with Trp 337 and strong cation- π interactions with Arg 234 (Fig. 3b, c). It is unusual to have such a high density of cation- π interactions within a protein structure²¹ (Fig. 3c).

Ultraviolet-B chromophore identification

Because UVR8 contains no external cofactor, the chromophore for ultraviolet-B perception must be amino acid(s). Among the 20 naturally occurring amino acids, only tryptophan and tyrosine, with maximal absorption wavelengths of 280 and 275 nm, respectively, are potentially capable of perceiving ultraviolet-B (280–315 nm). The UVR8 core domain contains thirteen tryptophan and eight tyrosine residues. Six of the thirteen tryptophans are located in the hydrophobic core and away from the homodimeric interface (Supplementary Fig. 7), making them unlikely candidates for ultraviolet-B perception. Among the eight tyrosine residues, only Tyr 253 is located close to the homodimeric interface (Fig. 3b). This analysis suggests that the chromophore for ultraviolet-B is among the seven tryptophan residues at the homodimeric interface.

Ultraviolet-B perception is probably coupled with chemical and/or conformational changes around the chromophore, presumably leading to alteration of its fluorescence emission spectra. Mutation of the

key chromophore should result in abrogation of its ability to sense ultraviolet-B and consequent loss of fluorescence changes that are characteristic of ultraviolet-B perception. To identify conclusively the chromophore, we individually mutated the seven tryptophan residues to phenylalanine.

First, we examined the intrinsic tryptophan fluorescence emission spectra for wild-type UVR8. In the absence of prior ultraviolet-B treatment, wild-type UVR8 was continuously monitored at an emission wavelength of 335 nm, with an ultraviolet-B excitation wavelength of 295 nm. Consistent with saturable ultraviolet-B perception, the fluorescence signal increased rapidly over the initial 200 s and reached a maximum after 400 s (Fig. 4a, top). Prolonged irradiation, however, resulted in a gradual decrease of the fluorescence signal (Supplementary Fig. 8a), presumably due to fluorescence quenching. With prior ultraviolet-B treatment, the fluorescence signal of wild-type UVR8 decreased very slowly over time (Fig. 4a, top), again due to fluorescence quenching. By sharp contrast, the UVR8 variants W285F and W285A completely lost the ability to perceive ultraviolet-B (Fig. 4a, bottom; Supplementary Fig. 8b). These observations identify Trp 285 as an essential component of the chromophore.

Next, we examined the other six UVR8 missense proteins. Most notably, the variant W233F no longer showed an increase in fluorescence induced by ultraviolet-B (Fig. 4b, left), suggesting that Trp 233 also has an essential role in ultraviolet-B perception. By sharp contrast, each of the five variants W337F, W302F, W250F, W198F and W94F retained the ability to sense ultraviolet-B and to undergo ultraviolet-B-induced dimer-to-monomer switches (Fig. 4b and Supplementary Fig. 8c; W94F, data not shown). These analyses unambiguously identify Trp 285 and Trp 233 collectively as the chromophore for ultraviolet-B perception. Consistent with this conclusion, Trp 285 and Trp 233 are involved in considerably more and stronger cation- π

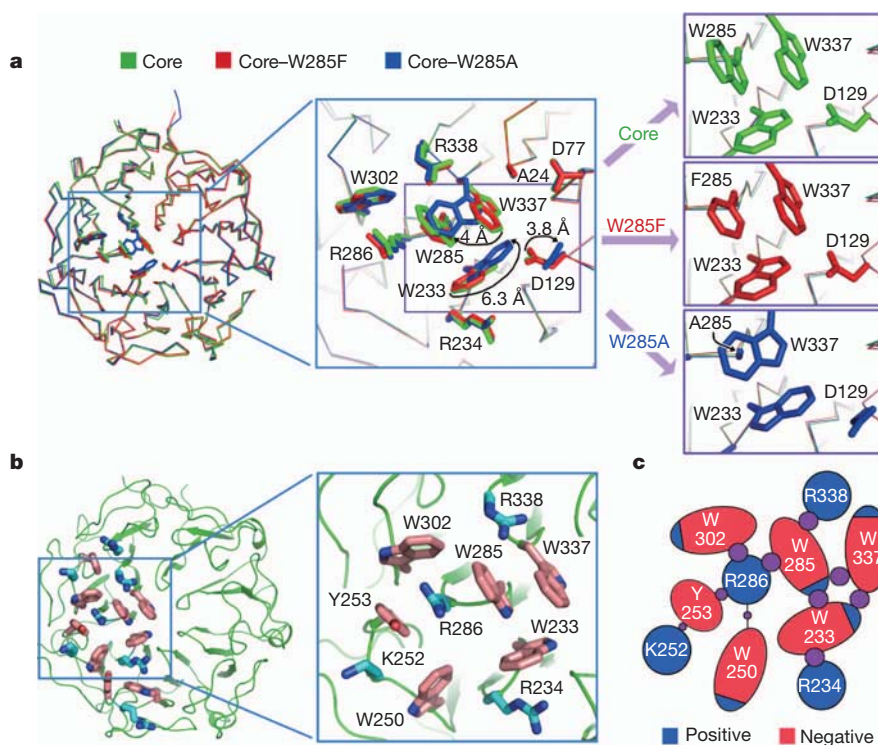


Figure 3 | Structural features of the ultraviolet-B-sensing amino acids. **a**, Structural comparison of UVR8 core domain (coloured green) with the variants W285F (red) and W285A (blue). Significant conformational changes occur to Trp 337, Trp 233 and Asp 129. **b**, A close-up view on the putative ultraviolet-B-sensing residues. Arginine and tryptophan residues are differentially coloured. **c**, A schematic diagram of the extensive network of

interactions compared to other tryptophans such as Trp 337 and Trp 198 (Supplementary Fig. 9).

Ultraviolet-B perception by Trp 285 and Trp 233 probably results in disruption of the cation- π interactions with key arginine residues, causing disruption of arginine-mediated intermolecular hydrogen bonds and consequent dissociation of UVR8 homodimers. To identify the key arginine residues, we generated five UVR8 variants, each targeting an arginine at the UVR8 homodimeric interface. Compared to wild-type UVR8, the variants R286A and R338A exhibited a grossly altered ability to perceive ultraviolet-B (Supplementary Fig. 10, left). Notably, these two UVR8 variants existed exclusively in a monomeric state (Supplementary Fig. 10, right), consistent with the important role of Arg 286 and Arg 338 at the homodimeric interface. By contrast, the variants R354A, R200A and R146A retained the abilities to perceive ultraviolet-B and to undergo an ultraviolet-B-induced dimer-to-monomer switch (Supplementary Fig. 10).

Transient nature of monomeric UVR8

Absorption of ultraviolet-B is predicted to excite the indole rings of the chromophore Trp 285 and Trp 233. Because the excited indole rings may dissipate energy to return to their ground state, we suspected that the ultraviolet-B-irradiated UVR8 monomers may revert back to homodimers over time. To examine this prediction, we subjected the UVR8 core domain to a saturating amount of ultraviolet-B radiation, left the protein at room temperature (23 °C) to recover in the absence of ultraviolet-B, and applied aliquots of the protein to SDS-PAGE at various time points. The result clearly shows that the UVR8 monomer slowly but steadily converted back to the homodimeric state (Fig. 4c). The partially recovered UVR8 homodimers can be completely monomerized again by ultraviolet-B irradiation (Fig. 4c, lane 8). The transient nature of monomeric UVR8 has obstructed repeated crystallization efforts, which were designed to capture the conformation of the ultraviolet-B-activated UVR8 core domain. The crystals were eventually

obtained but only contained the homodimeric form (Fig. 4d). Obviously, ultraviolet-B-irradiated UVR8 core domain in the crystallization drops slowly reverted to a dimeric state before crystallization. Nonetheless, we solved this structure, which is identical to that of the UVR8 core domain before ultraviolet-B irradiation.

Perspective

Photoreceptors rely on chromophores to perceive light. For example, phytochrome covalently associates with a single bilin molecule^{25,26}, cryptochrome contains a non-covalently-bound FAD molecule^{27,28}, and phototropin depends on FMN^{29,30}. In contrast to all known photoreceptors, UVR8 does not contain any external cofactor and instead uses two tryptophan residues, Trp 285 and Trp 233, as the chromophore for ultraviolet-B perception. This seems natural, because the absorption wavelengths for tryptophan coincide with the range of ultraviolet-B. Consequently, the ultraviolet-B-sensing mechanism of UVR8 differs markedly from those of other photoreceptors.

Our experimental evidence, in conjunction with knowledge of tryptophan fluorescence, yields a mechanistic model of ultraviolet-B perception by UVR8 (Fig. 4e). Ultraviolet-B irradiation results in excitation of the Trp 285 and Trp 233 indole rings, which is thought to disrupt the π -bond over the indole rings, leading to destabilization and abrogation of the intramolecular cation- π interactions. Such disruption is likely to trigger conformational switch of the side chains of Arg 286 and Arg 338, which would no longer be able to maintain intermolecular hydrogen bonds with Asp or Glu residues from the neighbouring UVR8 molecule, causing dissociation of the UVR8 homodimer (Fig. 4e). Furthermore, the excited indole rings are known to undergo a process of excited-state proton transfer³¹, which allows the indole ring to carry a positive charge and completely destroys the cation- π interactions. Importantly, excited-state proton transfer also leads to quenching of intrinsic tryptophan fluorescence, providing a plausible explanation for the observed slow decrease of

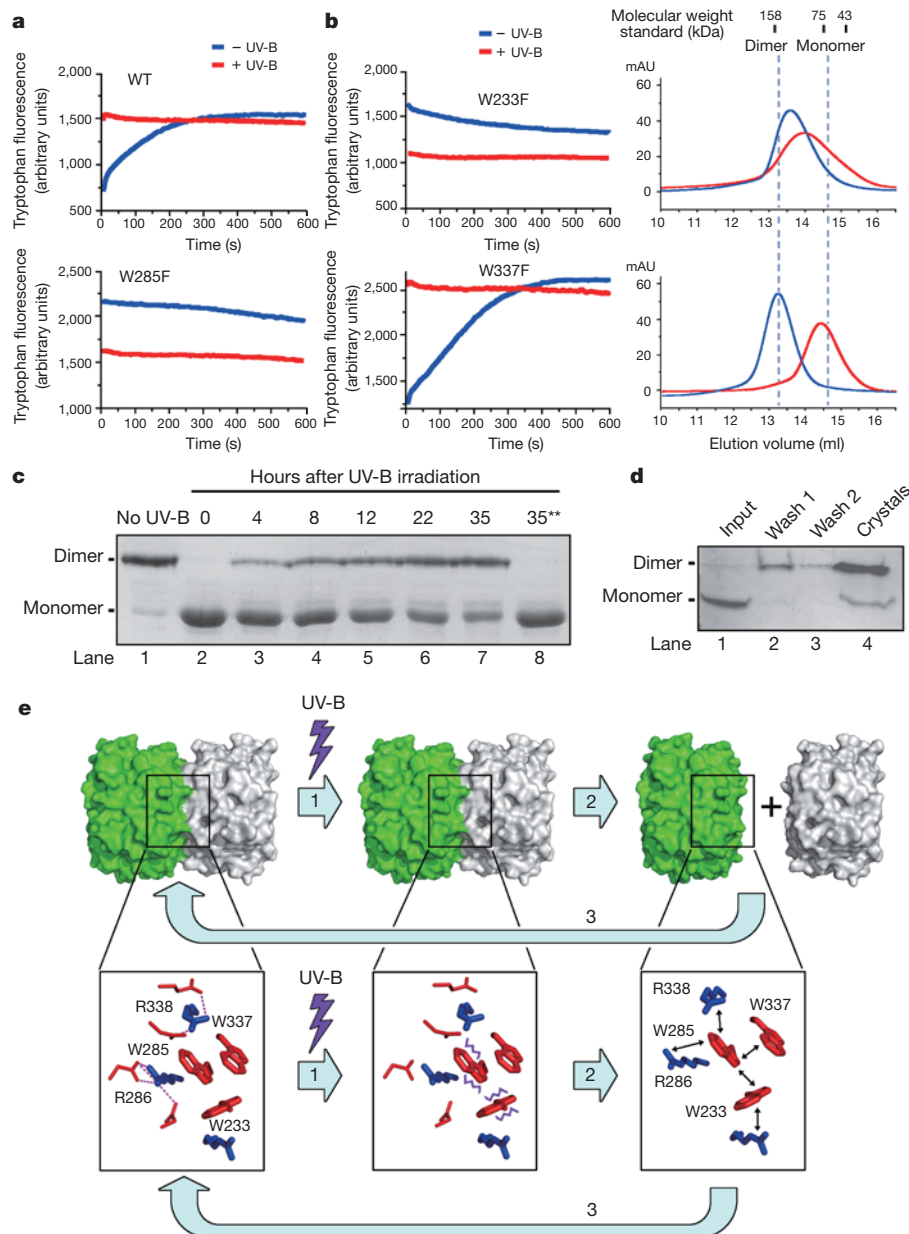


Figure 4 | Identification of Trp 285 and Trp 233 as the ultraviolet-B chromophore. **a**, Identification of Trp 285 as an essential ultraviolet-sensing amino acid. In the absence of prior ultraviolet-B (UV-B) radiation, the wild-type (WT) UVR8, but not the variant W285A, displayed a temporal, ultraviolet-B-induced increase of tryptophan fluorescence. This fluorescence increase completely disappeared after pre-irradiation by ultraviolet-B. **b**, The mutation W233F, but not W337F, in UVR8 led to loss of ultraviolet responsiveness as judged by tryptophan fluorescence (left panels). The variant W233F is no longer a stable homodimer, both with and without ultraviolet-B irradiation (top right). By contrast, the variant W337F can still undergo dimer-to-monomer switch in response to ultraviolet-B treatment (bottom right). **c**, Ultraviolet-B-irradiated UVR8 core domain slowly reverted back to a homodimeric state as judged by SDS-PAGE. Notably, 35 h after ultraviolet-B

fluorescence signal. Notably, Asp 129, Glu 182 and Arg 234 are all located in close proximity to Trp 233 and Trp 285 (Supplementary Fig. 11) and may well serve as proton donors. In this model, ultraviolet-B perception involves no covalent modification of UVR8, such as tryptophan oxidation or crosslinking. This notion is supported by mass spectrometric analysis (Supplementary Fig. 12). Further supporting this conclusion, ultraviolet-B-induced monomerization of UVR8 was unaffected by the presence of strong reducing or oxidizing agents (Supplementary Fig. 13).

irradiation, the partially recovered UVR8 homodimers were completely monomerized again by ultraviolet-B irradiation (lane 8, indicated by double asterisk). **d**, The crystals derived from ultraviolet-B-irradiated UVR8 core domain mostly contained the homodimeric form. Crystals were washed twice by crystallization buffer and examined by SDS-PAGE. **e**, A proposed mechanism for ultraviolet-B sensing by UVR8. In step 1, ultraviolet-B radiation excites Trp 285 and Trp 233. The excited states (purple wavy lines) of Trp 285 and Trp 233 are unable to maintain the cation- π interactions with surrounding residues. In step 2, disruption of the intramolecular cation- π interactions results in pronounced changes of side-chain conformations (black arrows), disrupting the intermolecular hydrogen bonds and causing dissociation of UVR8 homodimers. In step 3, Trp 285 and Trp 233 dissipate energy to return to the ground state, which allows re-formation of homodimers.

The observed fluorescence emission represents the total input from all 13 tryptophan residues and potentially other aromatic amino acids in UVR8. However, only those residues that are affected by ultraviolet-B perception or ultraviolet-B-induced environmental changes contribute to alteration of the fluorescence signal. The fact that a missense mutation of either Trp 285 or Trp 233 completely abrogates increase of the fluorescence signal identifies these two amino acids collectively as the chromophore for ultraviolet-B. In addition, this finding also strongly suggests that other tryptophan residues in UVR8 do not undergo any

significant ultraviolet-B-induced environmental changes. This analysis also validates the measurement of intrinsic tryptophan fluorescence as a sensitive method for the detection of ultraviolet-B perception.

Our study serves as a framework for understanding mutant phenotypes in plants. For example, the *Arabidopsis* mutants G145S and G202R lost the ability to perceive ultraviolet-B and the mutant proteins G145S and G202R were no longer able to form a homodimer¹². In our crystal structure, Arg 146 and Arg 200 make an important contribution to stabilize formation of the UVR8 homodimer (Fig. 2c); both Gly 145 and Gly 202 are located in close proximity to Arg 146 and Arg 200, respectively (Supplementary Fig. 14). Thus the mutations G145S and G202R may have a deleterious consequence on homodimer formation.

Despite revealing the structure revelation and underlying mechanism, important questions remain about the UVR8-mediated signalling pathway. It is unclear how UVR8 interacts with the central regulator of light signalling, COP1. How the UVR8–COP1 complex regulates downstream signalling events is uncertain. Answers to these questions await further experimental investigations.

While this manuscript was under final revision at *Nature*, Christie *et al.* reported similar findings in *Science*³².

METHODS SUMMARY

Wild-type UVR8 and all variants were subcloned by standard molecular biology, expressed, and purified. Tryptophan fluorescence was measured in a fluorescence spectrophotometer (HITACHI F-4600), with excitation and emission wavelengths of 295 and 335 nm, respectively. Gel filtration was used to observe ultraviolet-B-induced conformational changes. All crystals were generated by the hanging-drop vapour-diffusion method. All data sets were collected at the Shanghai Synchrotron Radiation Facility (SSRF) beamline BL17U and the SPring-8 beamline BL41XU and processed using the HKL2000 package³³. Further processing was carried out using programs from the CCP4 suite³⁴. The UVR8 structure was solved using the SeMet-derived UVR8(W285A) crystals. The selenium positions were determined using the program SHELXD³⁵. A partial model was traced automatically using the program BUCCANEER³⁶. The resulting map was in good quality and the partial model was manually rebuilt in COOT³⁷. Sequence assignment was aided with the selenium sites in the anomalous difference Fourier map. The final structure was refined with PHENIX³⁸. Using UVR8(W285A) coordinates as the initial search model, crystal structures of W285F and wild-type UVR8 were solved by molecular replacement using PHASER³⁹ and refined with COOT³⁷ and PHENIX³⁸.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 November 2011; accepted 10 February 2012.

Published online 29 February 2012.

- Falciatore, A. & Bowler, C. The evolution and function of blue and red light photoreceptors. *Curr. Top. Dev. Biol.* **68**, 317–350 (2005).
- van der Linden, A. M. *et al.* Genome-wide analysis of light- and temperature-entrained circadian transcripts in *Caenorhabditis elegans*. *PLoS Biol.* **8**, e1000503 (2010).
- Fogle, K. J., Parson, K. G., Dahm, N. A. & Holmes, T. C. CRYPTOCHROME is a blue-light sensor that regulates neuronal firing rate. *Science* **331**, 1409–1413 (2011).
- Jiao, Y., Lau, O. S. & Deng, X. W. Light-regulated transcriptional networks in higher plants. *Nature Rev. Genet.* **8**, 217–230 (2007).
- Kami, C., Lorrain, S., Hornitschek, P. & Fankhauser, C. Light-regulated plant growth and development. *Curr. Top. Dev. Biol.* **91**, 29–66 (2010).
- Quail, P. H. Phytochrome photosensory signalling networks. *Nature Rev. Mol. Cell Biol.* **3**, 85–93 (2002).
- Briggs, W. R. & Christie, J. M. Phototropins 1 and 2: versatile plant blue-light receptors. *Trends Plant Sci.* **7**, 204–210 (2002).
- Chaves, I. *et al.* The cryptochromes: blue light photoreceptors in plants and animals. *Annu. Rev. Plant Biol.* **62**, 335–364 (2011).
- Cashmore, A. R., Jarillo, J. A., Wu, Y. J. & Liu, D. Cryptochromes: blue light receptors for plants and animals. *Science* **284**, 760–765 (1999).
- Christie, J. M. Phototropin blue-light receptors. *Annu. Rev. Plant Biol.* **58**, 21–45 (2007).
- Liu, H., Liu, B., Zhao, C., Pepper, M. & Lin, C. The action mechanisms of plant cryptochromes. *Trends Plant Sci.* **16**, 684–691 (2011).
- Rizzini, L. *et al.* Perception of UV-B by the *Arabidopsis* UVR8 protein. *Science* **332**, 103–106 (2011).
- Möglich, A., Yang, X., Ayers, R. A. & Moffat, K. Structure and function of plant photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21–47 (2010).
- Kliebenstein, D. J., Lim, J. E., Landry, L. G. & Last, R. L. *Arabidopsis* UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity

- to human regulator of chromatin condensation 1. *Plant Physiol.* **130**, 234–243 (2002).
- Brown, B. A. *et al.* A UV-B-specific signaling component orchestrates plant UV protection. *Proc. Natl Acad. Sci. USA* **102**, 18225–18230 (2005).
- Brown, B. A. & Jenkins, G. I. UV-B signaling pathways with different fluence-rate response profiles are distinguished in mature *Arabidopsis* leaf tissue by requirement for UVR8, HY5, and HYH. *Plant Physiol.* **146**, 576–588 (2008).
- Kaiserli, E. & Jenkins, G. I. UV-B promotes rapid nuclear translocation of the *Arabidopsis* UV-B specific signaling component UVR8 and activates its function in the nucleus. *Plant Cell* **19**, 2662–2673 (2007).
- Li, D. & Roberts, R. WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell. Mol. Life Sci.* **58**, 2085–2097 (2001).
- Wall, M. A. *et al.* The structure of the G protein heterotrimer $G_{i\alpha 1}\beta_1\gamma_2$. *Cell* **83**, 1047–1058 (1995).
- Renault, L. *et al.* The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature* **392**, 97–101 (1998).
- Gallivan, J. P. & Dougherty, D. A. Cation– π interactions in structural biology. *Proc. Natl Acad. Sci. USA* **96**, 9459–9464 (1999).
- Gallivan, J. P. & Dougherty, D. A. A computational study of cation– π interactions vs salt bridges in aqueous media: implications for protein engineering. *J. Am. Chem. Soc.* **122**, 870–874 (2000).
- Dougherty, D. A. Cation– π interactions involving aromatic amino acids. *J. Nutr.* **137**, 1504S–1508S (2007).
- Sinnokrot, M. O., Valeev, E. F. & Sherrill, C. D. Estimates of the *ab initio* limit for π – π interactions: the benzene dimer. *J. Am. Chem. Soc.* **124**, 10887–10893 (2002).
- Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325–331 (2005).
- Uljasz, A. T. *et al.* Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature* **463**, 250–254 (2010).
- Brudler, R. *et al.* Identification of a new cryptochrome class. Structure, function, and evolution. *Mol. Cell* **11**, 59–67 (2003).
- Brautigam, C. A. *et al.* Structure of the photolyase-like domain of cryptochrome 1 from *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **101**, 12142–12147 (2004).
- Crosson, S. & Moffat, K. Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. *Proc. Natl Acad. Sci. USA* **98**, 2995–3000 (2001).
- Crosson, S. & Moffat, K. Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch. *Plant Cell* **14**, 1067–1075 (2002).
- Yu, H. T., Colucci, W. J., McLaughlin, M. L. & Barkley, M. D. Fluorescence quenching in indoles by excited-state proton transfer. *J. Am. Chem. Soc.* **114**, 8449–8454 (1992).
- Christie, J. M. *et al.* Plant UVR8 photoreceptor senses UV-B by tryptophan-mediated disruption of cross-dimer salt bridges. *Science* <http://dx.doi.org/10.1126/science.1218091> (9 February 2012).
- Otwiński, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computational Project. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
- Cowan, K. The Buccaneer software for automated model building. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
- DeLano, W. L. The PyMOL Molecular Graphics System. <http://www.pymol.org> (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. He and S. Huang at SSRF, and K. Hasegawa and T. Kumasaka at the SPring-8 beamline BL41XU, for assistance. This work was supported by funds from the Ministry of Science and Technology (grant no. 2009CB918801 to Y.S., and 2012CB910900 to X.W.D.), the National Natural Science Foundation, and the Beijing Municipal Commissions of Education and Science and Technology.

Author Contributions D.W., Q.H., H.D., X.W.D. and Y.S. designed all experiments. D.W., Q.H., Z.Y., W.C., C.Y. and J.Z. performed the experiments. D.W., Q.H., Z.Y., W.C., C.Y., X.H., J.Z., P.Y., H.D., J.W., X.W.D. and Y.S. contributed to technical work and data analysis. D.W., Q.H., Z.Y., W.C., C.Y., J.W., X.W.D. and Y.S. contributed to manuscript preparation. Y.S. wrote the manuscript.

Author Information The atomic coordinates and structure factor files of UVR8 wild type, W285A and W285F have been deposited in the Protein Data Bank under accession codes 4DNW, 4DNU and 4DNV, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.S. (shi-lab@tsinghua.edu.cn).

METHODS

Protein preparation. All constructs of *A. thaliana* UVR8 (full-length, residues 1–440) were cloned into pET-29b (Novagen) with a hexahistidine (6×His) tag at the C terminus. The plasmids were then transformed into *Escherichia coli* BL21(DE3) and overexpressed by induction with 0.2 mM β-D-thiogalactopyranoside (IPTG) at 18 °C overnight. The bacteria were harvested by centrifugation and resuspended in 150 mM NaCl, 25 mM Tris (pH 8.0), and lysed by sonication. After centrifugation, the supernatant was loaded into Ni²⁺-NTA affinity columns (Qiagen), and washed with 150 mM NaCl, 25 mM Tris (pH 8.0). The target protein was eluted by 250 mM imidazole (pH 8.0), 25 mM Tris (pH 8.0) and further purified by a Source-15Q column (GE Healthcare). The protein was then concentrated, and purified by gel filtration (Superdex-200, 10/30, GE Healthcare) with a buffer containing 150 mM NaCl, 25 mM Tris (pH 8.0) and 5 mM DTT. The proteins were ready for biochemical assay. The core domain of UVR8 (residues 12–381) was purified by gel filtration with the same protocol described above and concentrated to about 4 mg ml^{−1} for crystallization.

Ultraviolet-B radiation. The protein was exposed to ultraviolet-B radiation from an ultraviolet-B lamp (11 W, λ_{max} = 308 nm), at a distance of 20 cm, for 30 min on ice before biochemical assay or crystallization.

Tryptophan fluorescence measurements. The protein was adjusted to about 1 μM in the buffer containing 150 mM NaCl, 25 mM Tris (pH 8.0), 5 mM DTT. Tryptophan fluorescence was measured in a fluorescence spectrophotometer (HITACHI F-4600). The excitation and emission wavelengths were 295 and 335 nm, respectively.

Gel filtration. Superdex-200 (10/30, GE Healthcare) was used to observe ultraviolet-B-induced conformational changes. The column was pre-equilibrated with 150 mM NaCl, 25 mM Tris (pH 8.0), 5 mM DTT, and calibrated with molecular weight standards (GE Healthcare). The protein with or without ultraviolet-B irradiation was injected into the column, and eluted with a flow rate of 0.4 ml min^{−1}.

Mass spectrometry. The untreated and ultraviolet-irradiated wild-type, full-length UVR8 samples were mixed with 0.5 μl of 10 mg ml^{−1} α-cyano-4-hydroxysuccinamic acid in 50% acetonitrile, 0.1% (v/v) TFA, and applied onto a MALDI plate. MALDI mass spectra were recorded with a MALDI_TOF/TOF mass spectrometer operated in the linear mode. Bovine serum albumin (BSA) was used as the internal standard. The mass difference for the whole protein before and after ultraviolet-B treatment is less than 1 Da, indicating there is no mass change upon ultraviolet-B irradiation.

For detection of potential protein oxidation, the untreated and ultraviolet-irradiated UVR8 samples were separated on SDS-PAGE, excised, and in-gel digested with trypsin at 37 °C overnight. The peptides were extracted twice with 1% TFA in 50% acetonitrile for 30 min, and applied to LC-MS/MS analysis in the LTQ-Orbitrap mass spectrometer. The MS/MS spectra from each run were searched for possible tryptophan hydroxylation and formation of N-formyl-kyneurine. No apparent oxidation was detected for any tryptic fragment. This result suggests that ultraviolet radiation does not induce noticeable tryptophan oxidation.

Crystallization. SeMet-derived UVR8(W285A) crystals were grown at 18 °C using the hanging-drop vapour-diffusion method by mixing 1.2 μl of SeMet-derived UVR8(W285A) protein with 1.2 μl of reservoir solution containing 23% (w/v) PEG3350, 100 mM Bis-Tris buffer (pH 6.0) and 0.2 μl 30% (w/v) 1,5-diaminopentane dihydrochloride. Wild-type UVR8 crystals were obtained by mixing 1.5 μl of protein with an equal volume of reservoir solution containing 18% (w/v) PEG8000, 100 mM Tris buffer (pH 9.2) and 200 mM magnesium chloride. The UVR8 variant W285F was crystallized similarly using a reservoir solution containing 17% (w/v) PEG8000, 100 mM Tris buffer (pH 8.5) and 200 mM magnesium chloride. All native and SeMet crystals were directly flash-frozen in a cold nitrogen stream at 100 K.

Data collection and structural determination. The data sets for UVR8(W285F) and SeMet-derived UVR8(W285A) were collected at the SSRF beamline BL17U, the wild-type UVR8 data were collected at the SPring-8 beamline BL41XU. All data sets were integrated and scaled using the HKL2000 package³³. Further processing was carried out using programs from the CCP4 suite³⁴. Data collection statistics are summarized in Supplementary Table 1. The UVR8 structure was solved using the SeMet-derived UVR8(W285A) crystals in the *I*222 space group. The selenium positions were determined using the program SHELXD³⁵. The identified selenium positions were refined and initial phases were calculated using the PHASER SAD experimental phasing module³⁹. Solvent flattening and histogram matching were applied to the electron density map in DM³⁴. A crude partial model was traced automatically using the program BUCCANEER³⁶, then the model was fed back to the program PHASER to combine SAD phasing and partial structure information. The resulting map was in good quality and the partial model was manually rebuilt in COOT³⁷. Sequence assignment was aided with the selenium sites in the anomalous difference Fourier map. The final structure was refined with PHENIX³⁸. Using UVR8(W285A) coordinates as an initial model, crystal structures of W285F and wild-type UVR8 were solved by molecular replacement using PHASER³⁹ and manually refined with COOT³⁷ and PHENIX³⁸.

A close halo of large transparent grains around extreme red giant stars

Barnaby R. M. Norris¹, Peter G. Tuthill¹, Michael J. Ireland^{1,2,3}, Sylvestre Lacour⁴, Albert A. Zijlstra⁵, Foteini Lykou⁵, Thomas M. Evans^{1,6}, Paul Stewart¹ & Timothy R. Bedding¹

An intermediate-mass star ends its life by ejecting the bulk of its envelope in a slow, dense wind^{1–3}. Stellar pulsations are thought to elevate gas to an altitude cool enough for the condensation of dust¹, which is then accelerated by radiation pressure, entraining the gas and driving the wind^{2,4,5}. Explaining the amount of mass loss, however, has been a problem because of the difficulty of observing tenuous gas and dust only tens of milliarcseconds from the star. For this reason, there is no consensus on the way sufficient momentum is transferred from the light from the star to the outflow. Here we report spatially resolved, multiwavelength observations of circumstellar dust shells of three stars on the asymptotic giant branch of the Hertzsprung–Russell diagram. When imaged in scattered light, dust shells were found at remarkably small radii (less than about two stellar radii) and with unexpectedly large grains (about 300 nanometres in radius). This proximity to the photosphere argues for dust species that are transparent to the light from the star and, therefore, resistant to sublimation by the intense radiation field. Although transparency usually implies insufficient radiative pressure to drive a wind^{6,7}, the radiation field can accelerate these large grains through photon scattering rather

than absorption⁸—a plausible mass loss mechanism for lower-amplitude pulsating stars.

We observed W Hydrae, R Doradus and R Leonis using aperture-masked^{9,10} polarimetric interferometry (Fig. 1), along with dust-free stars to verify our detection methodology. Figure 1 shows the ratio of the horizontally and vertically polarized visibilities ($V_{\text{horiz}}/V_{\text{vert}}$), plotted as a function of baseline azimuth and length. The dust-free star 2 Centauri, which has no polarized flux, shows a constant ratio $V_{\text{horiz}}/V_{\text{vert}} = 1.0$ within its uncertainties. However, the dust-enshrouded star W Hya shows a strong sinusoidal variation of $V_{\text{horiz}}/V_{\text{vert}}$ with azimuth, as expected from a resolved circumstellar scattering shell. By taking advantage of spherical symmetry, we produced the much simpler baseline-dependent observable plotted in Fig. 2. A model was then fitted to the data to determine the dust shell radius and the amount of light scattered by the shell at each wavelength (Supplementary Information). These results are summarized in Table 1. Figure 3 shows the model image of the star and shell as seen in orthogonal polarizations for W Hya at a wavelength of 1.24 μm , from which the model visibilities were derived.

Scattered-light dust shells around the three asymptotic giant branch (AGB) stars observed were found close to the star, at radii $\lesssim 2$ stellar radii. This is in contrast to earlier models that place the shell at many stellar radii², but is consistent with some recent models⁶ and with interferometric¹¹ and polarimetric¹² measurements. On the basis of typical elemental abundances and spectral observations, the composition of

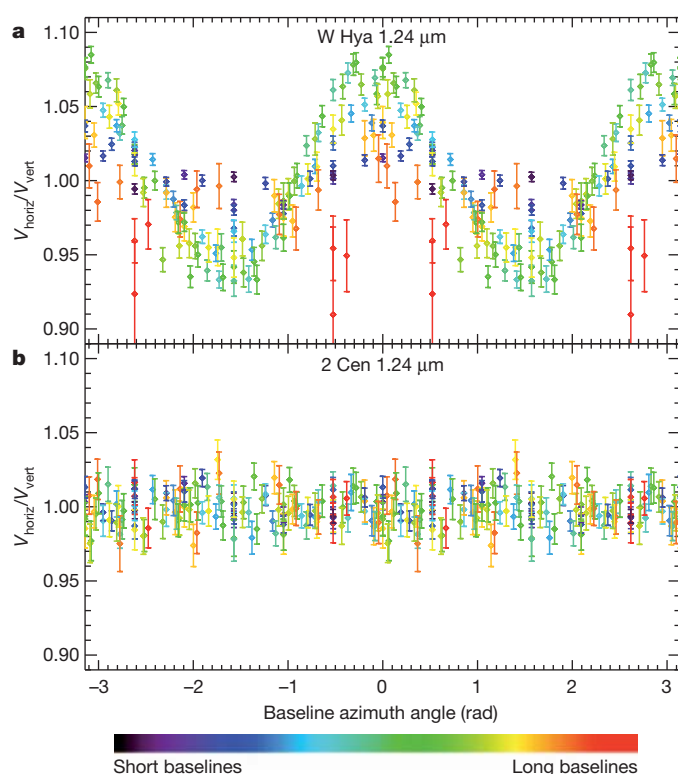


Figure 1 | Polarimetric interferometry of W Hya at 1.24 μm . Although light scattered by each part of the circumstellar dust shell is strongly polarized, the integrated signal recovered with conventional polarimetry is zero for an unresolved spherically symmetric shell. In this study, aperture-masking interferometry^{9,10} (which converts the 8-m pupil of the Very Large Telescope into a multi-element interferometer, using the NACO²⁴ instrument) allows access to the ~ 10 -mas spatial scales required to resolve the shell, and polarimetric measurements (obtained by simultaneously measuring interferometric visibilities in orthogonal polarizations²⁵) allows light from the star and light from the dust shell to be differentiated. Here the ratio of the horizontally polarized visibilities (V_{horiz}) to vertically polarized visibilities (V_{vert}) is plotted against baseline azimuth angle (corresponding to position angle on the sky). Colour encodes the baseline length (longest, 7.3 m; shortest, 0.56 m). The ratio $V_{\text{horiz}}/V_{\text{vert}}$ is a differential observable, which allows the cancellation of residual systematic errors and depends only on the fractional polarized scattered light signal. **a**, Result for W Hya, an AGB star with a circumstellar shell; $V_{\text{horiz}}/V_{\text{vert}}$ deviates from 1, varying sinusoidally with azimuth. This is the signal expected from a thin, spherically symmetric dust shell scattering the light from a central star. This signal varies in amplitude for different baselines, encoding the spatial extent of the resolved structure. Data points have been repeated over two cycles. The longest baselines (red) have poor signal-to-noise ratios because they are close to the null, where the visibility curve of the star is extremely low. Error, 1σ . **b**, Visibility data for the star 2 Cen, which has no circumstellar dust shell and, hence, no polarized signal from scattering; here $V_{\text{horiz}}/V_{\text{vert}} \approx 1$ for all azimuths. Errors, 1σ .

¹Sydney Institute for Astronomy, School of Physics, University of Sydney, New South Wales 2006, Australia. ²Department of Physics and Astronomy, Macquarie University, New South Wales 2109, Australia. ³Australian Astronomical Observatory, PO Box 296, Epping, New South Wales 1710, Australia. ⁴LESIA-Observatoire de Paris, CNRS, Université Pierre et Marie Curie, Université Paris-Diderot, Meudon 92190, France. ⁵Jodrell Bank Centre for Astrophysics, School of Physics & Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK. ⁶Department of Physics, Denys Wilkinson Building, Keble Road, University of Oxford, Oxford OX1 3RH, UK.

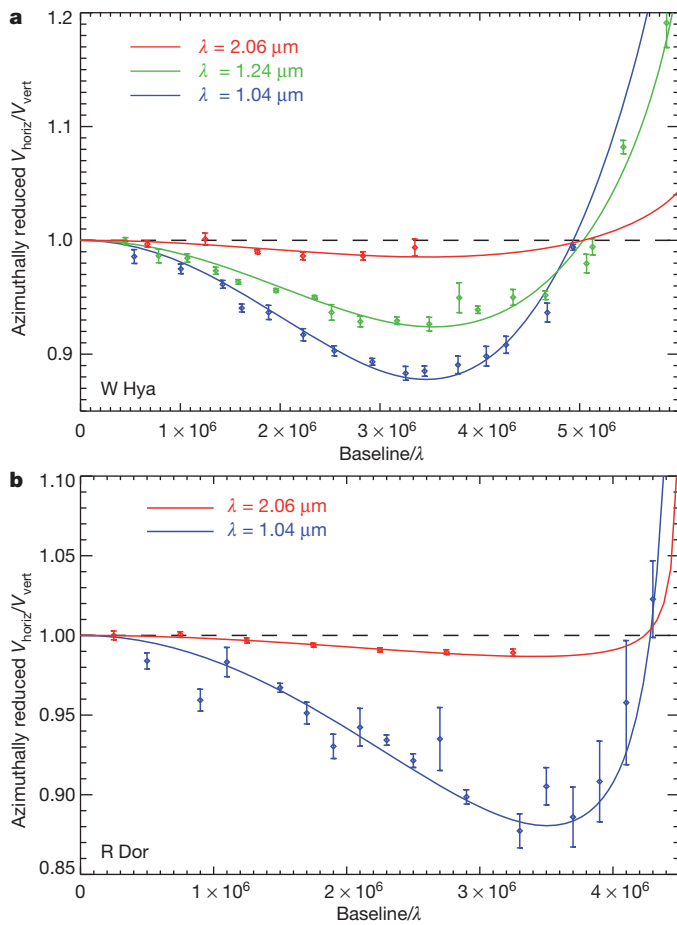


Figure 2 | Wavelength dependence of scattering for W Hya and R Dor. The azimuthally reduced ratio $V_{\text{horiz}}/V_{\text{vert}}$ for the AGB stars W Hya and R Dor, plotted against spatial frequency (baseline length divided by wavelength, λ), at multiple wavelengths. The functional form of the visibility ratio excursions of all three AGB stars are consistent, within uncertainties, with a simple spherically symmetric shell. We were therefore able to enhance our signal-to-noise ratio significantly in a quantitative analysis by reducing our two-dimensional visibility data to a one-dimensional function of the baseline length (corresponding to spatial frequency). This was achieved by dividing $V_{\text{horiz}}/V_{\text{vert}}$ by the expected sinusoidal variation (characteristic of a spherical shell) at a fixed amplitude, resulting in the much simpler, baseline-dependent observable plotted here. Points are the observed data (binned; errors, 1σ) and the solid lines are the fitted model. The spatial frequency of the minimum of the characteristic ‘dip’ varies with the radius of the dust shell: for a larger shell, the minimum of the dip occurs at lower spatial frequencies. The depth of the dip depends on the amount of scattered light, with a larger deviation from $V_{\text{horiz}}/V_{\text{vert}} = 1.0$ indicating that a larger fraction of the total flux arises from light scattered by the shell. This is seen to decrease strongly at longer wavelengths as expected theoretically; the precise change in this quantity as a function of wavelength can be used to determine the dust grain radius using Mie scattering theory (Fig. 4). The fitted parameters for these quantities are included in Table 1. The 2.06- μm data have insufficient spatial resolution to constrain the shell size, so for models at these wavelengths we fixed the shell size to be consistent with the fitted size at shorter wavelengths.

possible solution to this dilemma arises when very large grains are considered.

The degree of scattering by dust depends strongly on the wavelength of the incident light and on the size of the particles. By analysing our multiwavelength measurements using Mie scattering theory, we determined the effective grain size and the number of grains. As shown in Fig. 4, we found an effective grain radius of ~ 300 nm. For grains of this size, the scattering opacity becomes very large, well beyond that resulting from Rayleigh scattering when the particles are smaller. In this regime, the contribution to radiative acceleration by scattering, rather than by absorption alone, must be considered. Models show that grains exceeding a certain critical scattering opacity can drive a wind at high magnesium condensation and that, for a star of temperature 2,700 K, this critical scattering opacity is only exceeded in a narrow range of dust grain radii around 300 nm (ref. 8). We also note that a narrow range of grain radii, of the order of ~ 500 nm, is predicted on the basis of a self-regulating feedback mechanism: grain growth effectively halts once the critical size is reached, because the dust is then accelerated outwards and gas densities quickly decrease⁸. This is consistent with observations of grains in the interstellar medium, which are dominated by silicates¹⁵ and have similar grain sizes^{16,17}. Wind driving due to scattering by magnesium-rich silicates is consistent with the finding that mass loss in AGB stars depends on their metallicity¹⁸. Although this model encounters difficulties in the case of stars with extremely extended atmospheres, such as R Leo (owing to the mass of the stellar atmosphere at and above the dust-forming layers being too high to allow sufficient acceleration¹⁹), it provides a plausible explanation for the mass loss of semiregular pulsating stars such as R Dor. Our

AGB dust shells is expected to be dominated by silicates^{5,13,14} in the form of olivine ($\text{Mg}_{2x}\text{Fe}_{2(1-x)}\text{SiO}_4$) and/or pyroxene ($\text{Mg}_x\text{Fe}_{(1-x)}\text{SiO}_3$), where $0 \leq x \leq 1$. The temperature of a grain is determined by its opacity, that is, how strongly it absorbs the surrounding radiation field. Multiwavelength models⁶ show that silicate dust that contains iron absorbs the stellar flux strongly (as these dust species have high opacities at wavelengths of $\sim 1 \mu\text{m}$, where the energy distribution peaks) and so can only condense at distances greater than ~ 5 stellar radii. These iron-rich species could be accelerated by absorption of stellar radiation, but they form too far from the star to provide an efficient mass loss mechanism for low-amplitude pulsators⁶ (semiregular variable stars). Our detection of dust much closer to the star is instead consistent with the presence of iron-free silicates such as forsterite (Mg_2SiO_4) and enstatite (MgSiO_3), which are almost transparent at wavelengths of $\sim 1 \mu\text{m}$. Such grains do not heat to sublimation, despite the intense radiation close to the star, but the same transparency also prevents the momentum transfer from starlight required to drive a wind. For some stars, a

Table 1 | Summary of fitted model parameters

| Star | Φ | λ (μm) | R_{star} (mas) | R_{shell} (mas) | Scattered fraction | Grain radius (nm) | Scattering-shell mass |
|-------|--------|-----------------------------|-------------------------|--------------------------|--------------------|-------------------|--|
| R Dor | 0.7 | 1.04 | 27.2 ± 0.2 | 43.3 ± 0.3 | 0.124 ± 0.003 | 299 ± 39 | $(2.7 \pm 0.2) \times 10^{-10} M_{\odot}$ |
| | | 2.06 | 27.7 ± 1.4 | 43.6 ± 3.2 | 0.014 ± 0.002 | | |
| W Hya | 0.2 | 1.04 | 18.7 ± 0.4 | 37.9 ± 0.2 | 0.176 ± 0.002 | 316 ± 4 | $(1.04 \pm 0.02) \times 10^{-9} M_{\odot}$ |
| | | 1.24 | 18.9 ± 0.5 | 37.0 ± 0.3 | 0.110 ± 0.003 | | |
| | | 2.06 | 18.9 (fixed) | 37.0 (fixed) | 0.022 ± 0.004 | | |
| | | 1.04 | 18.3 ± 0.3 | 29.9 ± 0.4 | 0.120 ± 0.003 | | |
| R Leo | 0.4 | 1.04 | | | | $\sim 300^*$ | $\sim 2 \times 10^{-10} M_{\odot}$ |

The radii of the dust shells were found to be $\leq 2R_{\text{star}}$. The scattered fraction is the proportion of the total flux arising from scattering by the dust shell. The grain radius was obtained from fitting to multiwavelength observations using Mie scattering (with the value for R Leo fixed at 300 nm). The scattering-shell mass was calculated only for the observed dust (and not, for example, for a distribution extending to small grains invisible to our technique). Stellar radii given are for a uniform disc. All three AGB stars (W Hya, R Dor and R Leo) were observed in March 2009, with additional observations of W Hya at 1.24 μm and 2.06 μm made in June 2010. Although the photospheric diameter, and possibly the dust shell diameter, are expected to vary throughout the stellar pulsation cycle, the two sets of observations for W Hya were taken approximately one period apart and have therefore been combined. These figures assume both the dust to be iron-free silicate (forsterite) grains of uniform size and there to be full magnesium condensation. Full magnesium condensation is a reasonable assumption for the stars with more compact atmospheres (for example R Dor) but is inconsistent with observed optical depths for stars with more extended atmospheres¹⁹. In the event that there is also a population of small, weakly scattering grains that do not show up in our data, these values represent lower limits, with the total shell mass being greater. Furthermore, if the shell is geometrically extended then the true mass will be greater, as these calculations assume a thin shell. The uncertainties given are based on random errors and do not account for systematic errors such as those described here. Hipparcos parallaxes²² and experimentally measured optical constants²³ have been used. Φ indicates visual phase, derived from the AAVSO International Database; R_{star} , stellar radius; R_{shell} , radius of dust shell.

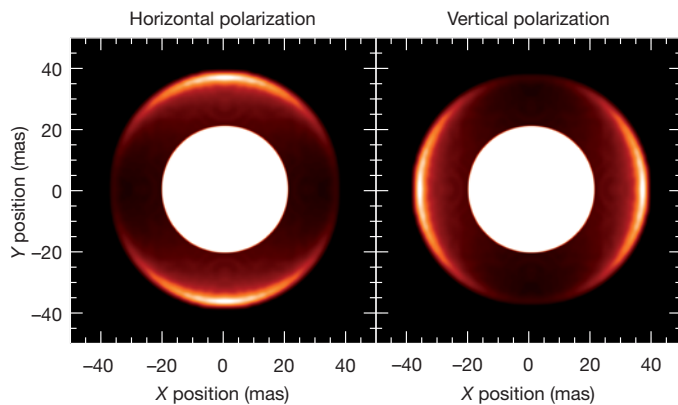


Figure 3 | Model image for W Hya with circumstellar shell viewed in horizontally and vertically polarized light. The white disc represents the uniform-disc star used in the model. A three-dimensional model of a star with a thin scattering shell was constructed, and the scattered intensity observed in each polarization for each point on the shell was calculated using Mie scattering, yielding an image of the star and shell. We then derived polarized visibilities from the model and fitted them to the observed visibilities, to determine the dust shell radius and the scattered fraction. Details of the modelling process can be found in Supplementary Information. See Supplementary Fig. 2 for a diagram illustrating how the polarized intensity distribution arises.

observations provide direct evidence for a population of dust grains capable of powering a scattering-driven wind.

The last column of Table 1 gives the mass of the dust that contributes to the observed scattering signal, assuming the shell to be thin and the dust grains to be forsterite of a uniform size. If full magnesium condensation and solar abundances are assumed, then the gas-to-dust ratio is ~ 600 , which yields total shell masses of $\sim 6 \times 10^{-7} M_{\odot}$, $\sim 2 \times 10^{-7} M_{\odot}$ and $\sim 1 \times 10^{-7} M_{\odot}$ for W Hya, R Dor and R Leo, respectively. Because the pulsation periods of these stars are ~ 1 yr and the mass loss rates are $\sim 1 \times 10^{-7} M_{\odot} \text{ yr}^{-1}$ (refs 20, 21), this implies that for stars with less extended atmospheres a large fraction of the observed shell is ejected each pulsation cycle, consistent with the observed dust being part of an outflow. In the extended-atmosphere case, where full magnesium condensation is not observationally supported, a possible alternative dust species is corundum (Al_2O_3),

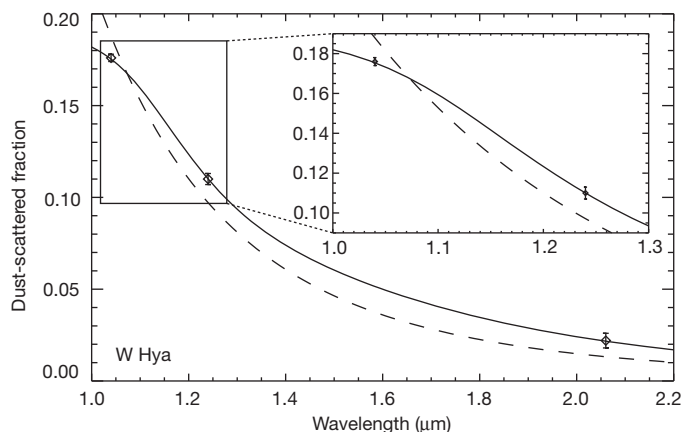


Figure 4 | Grain size measurement. Grain size fitted to the fraction of scattered light as a function of wavelength, for W Hya. Inset, magnified view of the 1.0–1.3- μm region. The solid line represents the fitted Mie scattering model (where grain size and grain number were fit parameters), and the dashed line represents the best Rayleigh fit (where grain size was fixed to be below the Rayleigh limit). The data are inconsistent with Rayleigh scattering. The fitted Mie model yields an effective grain radius of 316 ± 4 nm. In reality, a distribution of grain sizes may be present; for example, a population of very much smaller particles would contribute only weakly and could be undetected. Our data show that, regardless of the presence or absence of smaller grains, a population of large, ~ 300 -nm, grains is required. Errors, 1σ .

as discussed in Supplementary Information. Further time-dependent grain growth and dynamical models will help elucidate the role of light scattered from large-grained dust in the process of mass loss from AGB stars.

Received 25 October 2011; accepted 9 February 2012.

- Wood, P. R. Pulsation and mass loss in Mira variables. *Astrophys. J.* **227**, 220–231 (1979).
- Habing, H. J. Circumstellar envelopes and asymptotic giant branch stars. *Astron. Astrophys. Rev.* **7**, 97–207 (1996).
- Marengo, M. A review of AGB mass loss imaging techniques. *Proc. Astron. Soc. Aust.* **26**, 365–371 (2009).
- Gustafsson, B. & Höfner, S. *Asymptotic Giant Branch Stars* (eds Habing, H. J. & Olofsson, H.) Ch. 4 (Springer, 2004).
- Höfner, S. Dust formation and winds around evolved stars: the good, the bad and the ugly cases. *Astron. Soc. Pacif. Conf. Ser.* **414**, 3–21 (2009).
- Woitke, P. Too little radiation pressure on dust in the winds of oxygen-rich AGB stars. *Astron. Astrophys.* **460**, L9–L12 (2006).
- Ireland, M. J. & Scholz, M. Observable effects of dust formation in dynamic atmospheres of M-type Mira variables. *Mon. Not. R. Astron. Soc.* **367**, 1585–1593 (2006).
- Höfner, S. Winds of M-type AGB stars driven by micron-sized grains. *Astron. Astrophys.* **491**, L1–L4 (2008).
- Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H. & Haniff, C. A. Michelson interferometry with the Keck I Telescope. *Publ. Astron. Soc. Pacif.* **112**, 555–565 (2000).
- Lacour, S. *et al.* Sparse aperture masking at the VLT. I. Faint companion detection limits for the two debris disk stars HD 92945 and HD 141569. *Astron. Astrophys.* **532**, A72 (2011).
- Wittkowski, M. *et al.* The extended atmospheres of Mira variables probed by VLTI, VLBA, and APEX. *Astron. Soc. Pacif. Conf. Ser.* **445**, 107–112 (2011).
- Ireland, M. J., Tuthill, P. G., Davis, J. & Tango, W. Dust scattering in the Miras R Car and RR Sco resolved by optical interferometric polarimetry. *Mon. Not. R. Astron. Soc.* **361**, 337–344 (2005).
- Waters, L. B. F. M. *et al.* Mineralogy of oxygen-rich dust shells. *Astron. Astrophys.* **315**, L361–L364 (1996).
- Gail, H.-P. in *Astromineralogy* (ed Henning, T. K.) 55–120 (Lect. Notes Phys. 609, Springer, 2003).
- Draine, B. T. Interstellar dust grains. *Annu. Rev. Astron. Astrophys.* **41**, 241–289 (2003).
- Frisch, P. C. *et al.* Dust in the local interstellar wind. *Astrophys. J.* **525**, 492–516 (1999).
- Weingartner, J. C. & Draine, B. T. Dust grain-size distributions and extinction in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud. *Astrophys. J.* **548**, 296–309 (2001).
- Lagadec, E. & Zijlstra, A. A. The trigger of the asymptotic giant branch superwind: the importance of carbon. *Mon. Not. R. Astron. Soc.* **390**, L59–L63 (2008).
- Ireland, M. J., Scholz, M. & Wood, P. R. Dynamical opacity-sampling models of Mira variables - II. Time-dependent atmospheric structure and observable properties of four M-type model series. *Mon. Not. R. Astron. Soc.* **418**, 114–128 (2011).
- Loup, C., Forveille, T., Omont, A. & Paul, J. F. CO and HCN observations of circumstellar envelopes. A catalogue—mass loss rates and distributions. *Astron. Astrophys. Suppl. Ser.* **99**, 291–377 (1993).
- Olofsson, H., González Delgado, D., Kerschbaum, F. & Schöier, F. L. Mass loss rates of a sample of irregular and semiregular M-type AGB-variables. *Astron. Astrophys.* **391**, 1053–1067 (2002).
- van Leeuwen, F. Validation of the new Hipparcos reduction. *Astron. Astrophys.* **474**, 653–664 (2007).
- Jäger, C., Dorschner, J., Mutschke, H., Posch, T. & Henning, T. Steps toward interstellar silicate mineralogy. VII. Spectral properties and crystallization behaviour of magnesium silicates produced by the sol-gel method. *Astron. Astrophys.* **408**, 193–204 (2003).
- Lenzen, R. *et al.* NAOS-CONICA first on sky results in a variety of observing modes. *Proc. SPIE* **4841**, 944–952 (2003).
- Tuthill, P. *et al.* Sparse aperture masking (SAM) at NAOS/CONICA on the VLT. *Proc. SPIE* **7735**, 773510 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was based on observations collected with the NACO instrument at the European Southern Observatory, Chile.

Author Contributions P.G.T., M.J.I. and S.L. commissioned the SAMPol observing mode and conducted the initial observations, and developed the standard aperture-masking data reduction pipeline. B.R.M.N. developed the polarimetric data reduction and modelling/fitting procedures, analysed the data and conducted later observations. T.M.E. and P.S. contributed to the data analysis procedures. T.R.B., A.A.Z. and F.L. contributed to both the observing programme and the interpretation of the findings.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.R.M.N. (bnorris@physics.usyd.edu.au).

Layered boron nitride as a release layer for mechanical transfer of GaN-based devices

Yasuyuki Kobayashi¹, Kazuhide Kumakura¹, Tetsuya Akasaka¹ & Toshiaki Makimoto¹

Nitride semiconductors are the materials of choice for a variety of device applications, notably optoelectronics^{1,2} and high-frequency/high-power electronics³. One important practical goal is to realize such devices on large, flexible and affordable substrates, on which direct growth of nitride semiconductors of sufficient quality is problematic. Several techniques—such as laser lift-off^{4,5}—have been investigated to enable the transfer of nitride devices from one substrate to another, but existing methods still have some important disadvantages. Here we demonstrate that hexagonal boron nitride (h-BN) can form a release layer that enables the mechanical transfer of gallium nitride (GaN)-based device structures onto foreign substrates. The h-BN layer serves two purposes: it acts as a buffer layer for the growth of high-quality GaN-based semiconductors, and provides a shear plane that makes it straightforward to release the resulting devices. We illustrate the potential versatility of this approach by using h-BN-buffered sapphire substrates to grow an AlGaIn/GaN heterostructure with electron mobility of $1,100 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$, an InGaIn/GaN multiple-quantum-well structure, and a multiple-quantum-well light-emitting diode. These device structures, ranging in area from five millimetres square to two centimetres square, are then mechanically released from the sapphire substrates and successfully transferred onto other substrates.

State-of-the-art growth technologies, such as metal–organic vapour phase epitaxy (MOVPE) and molecular beam epitaxy, enable single-crystal high-quality nitride devices to be grown on sapphire, silicon carbide and silicon substrates, but not on polycrystalline or amorphous substrates⁶. Releasing single-crystal high-quality devices from one substrate and transferring them to another is one of the promising ways to address this limitation. In general, nitride semiconductors have been grown on sapphire substrates using a buffer layer: such a layer can consist of low-temperature AlN (ref. 7), low-temperature GaN (ref. 8) or AlON (ref. 9). However, releasing nitride semiconductors from sapphire substrates is difficult owing to the strong covalent sp^3 bonding between the buffer layer and the semiconductors. To overcome this difficulty, thermal release with laser radiation^{4,5}, stamp-based printing¹⁰, chemical release by etching of a sacrificial layer^{11–13}, wet chemical etching^{14–17}, and mechanical release¹⁸ from sapphire substrates have been investigated. Compared with the thermal and chemical approaches, mechanical release is, in principle, an uncomplicated technique requiring neither chemical treatments nor additional equipment for the transfer. However, damage, the limited size of the released GaN, and limited throughput in mechanical releases have remained serious issues. Graphite, graphene¹⁹ and h-BN act as ideal layers for releasing devices from substrates mechanically. However, direct growth of nitride semiconductors on graphene is impossible, and an intermediate layer, such as a zinc oxide (ZnO) nano-wall, is therefore needed¹⁸. In addition, the transfer of nitride semiconductors grown on graphene lacks scalability. On the other hand, we expected that high-quality nitride semiconductors could be grown on h-BN with an AlN or AlGaIn buffer with high scalability because h-BN is itself a nitride semiconductor and can be grown on a substrate uniformly. In an earlier study, polycrystalline BN films were used as buffer layers to

grow GaN films on (001) Si substrates; however, the grown GaN was polycrystalline²⁰. In addition, the study was restricted to the use of BN as a buffer layer. Here, as a new release layer, we use single-crystal h-BN, which not only acts as a buffer layer for a nitride device but also allows us to release the device from the substrate and transfer it to a foreign substrate mechanically in a highly scalable way. This versatile approach has the potential to be scaled up to production size (several inches) at low cost.

The single-crystal h-BN layer with an atomically flat surface enables us to grow an AlGaIn/GaN heterostructure and InGaIn-based multiple quantum well (MQW) and MQW LED (light-emitting diode) structures and to release these structures from the host sapphire substrates and transfer them to foreign substrates. Figure 1 depicts our materials design and the release and transfer processes. First, we grew a single-crystal (0001) h-BN ultrathin layer on a (0001) sapphire substrate^{21,22} (Fig. 1a). The orientation relationship between the substrate and the h-BN is $(0001)_{\text{h-BN}} \parallel (0001)_{\text{sapphire}}$, where the plane of boron and nitrogen bonded with sp^2 hybridization is parallel to the substrate surface. Then, the MQW structure—consisting of a 0.3- μm -thick AlGaIn layer, a 3- μm -thick GaN layer, a ten-period InGaIn/GaN MQW structure, and a 0.1- μm -thick GaN layer—was grown on the h-BN release layer (Fig. 1a). After the MOVPE growth, we flipped the MQW structure upside down and put it on a foreign substrate via an adhesive sheet (an indium sheet in this case) (Fig. 1b). Next, the MQW structure attached to the indium sheet on the foreign substrate was heated to a temperature sufficient to heat-seal the indium to the sapphire and the MQW. (Fig. 1c). Finally, the MQW structure was released from the host sapphire substrate by mechanical force and the MQW was thereby transferred to the foreign substrate (Fig. 1d). The mechanical force easily separates the MQW from the host sapphire and the separation occurs within the h-BN release layer owing to van der Waals forces of the h-BN layered structure. In this way, MQWs and other types of nitride devices can be transferred to all kinds of substrates, such as silicon, polycrystalline metal, glass and transparent plastics.

An AlN or AlGaIn layer on the h-BN layer makes it possible to greatly improve the surface morphology and crystalline quality of nitride semiconductors grown on it. GaN directly grown on the h-BN layer has a rough and irregular island-shaped surface morphology, and is polycrystalline (Supplementary Fig. 1). Hence, we first grew a wurtzite AlN layer on the h-BN and then grew a GaN film on the AlN to overcome the difficulty of direct growth of GaN (Fig. 2a). In contrast to GaN grown on the h-BN directly, GaN grown on the AlN shows a step-like flat surface with measured root mean square (r.m.s.) roughness of 0.69 nm over an area of $5 \times 5 \mu\text{m}^2$, as seen in the atomic force microscopy (AFM) image in Fig. 2b. X-ray diffraction in the $2\theta/\omega$ configuration exhibits only GaN (0002) and AlN (0002) diffraction peaks (Fig. 2c), indicating that (0001) single-crystal GaN film has been grown on the (0001) AlN on the h-BN layer. The orientation relationship between the GaN, AlN, h-BN and the sapphire substrate is described in Supplementary Figs 2 and 3. The weak beam dark-field transmission electron microscopy (TEM) images in Fig. 2d and e reveal that the AlN layer, working as a dislocation filter, decreases

¹NTT Basic Research Laboratories, Nippon Telegraph and Telephone Corporation 3-1, Morinosato Wakamiya, Atsugi, Kanagawa 243-0292, Japan.

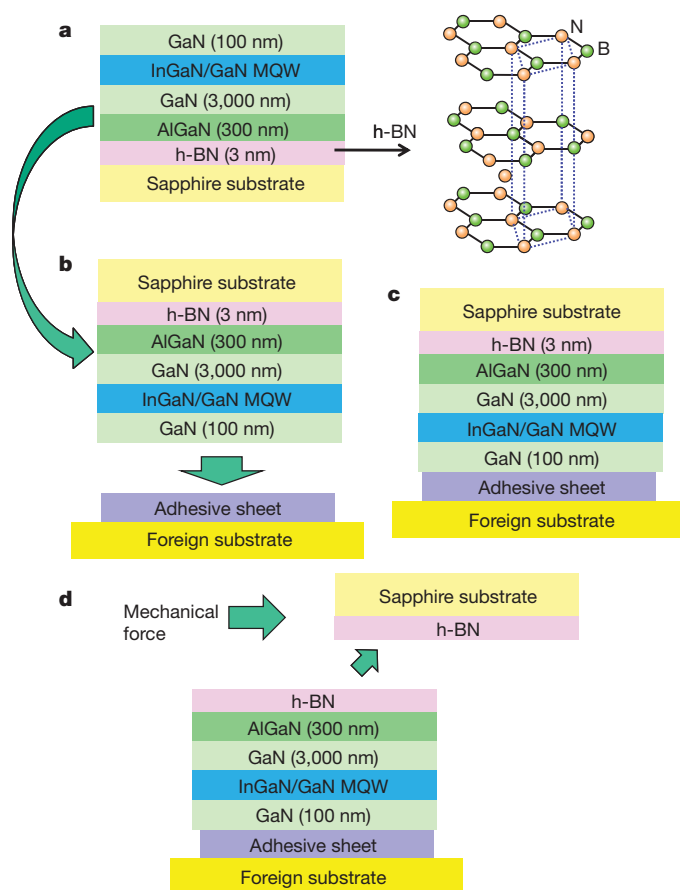


Figure 1 | Schematic illustrations of the MQW materials design, release and transfer processes. **a**, Single-crystal h-BN release layer growth is performed on a sapphire substrate, followed by growth of the wurtzite AlGaIn layer and the MQW structure. On the right, the crystal structure of h-BN is shown. The dotted lines mark the unit cell dimensions of h-BN. **b**, The MQW structure is flipped upside down and placed on a foreign substrate via an adhesive sheet (indium sheet). **c**, The MQW structure attached to the sheet is heated to a temperature sufficient to heat-seal the indium, and then the MQW structure is released from the host sapphire substrate by mechanical force and transferred to the foreign substrate (**d**).

the density of threading dislocations in the GaN and that the predominant defects in the GaN are mixed dislocations with a density of $8.6 \times 10^9 \text{ cm}^{-2}$ (Supplementary Information part III). In the GaN film, the tilt and twist angles were respectively evaluated to be 0.16° and 0.47° from the full-width at half-maximum (FWHM) of the X-ray rocking curve profiles obtained by fitting the rocking curve profiles to a pseudo-Voigt function^{23,24}. Figure 2f shows a convergent beam electron diffraction (CBED) pattern experimentally obtained for the GaN layer and a simulated CBED pattern for Ga-terminated GaN. The good agreement between these patterns reveals that the GaN film has a Ga termination (Fig. 2f). In addition, to verify the device quality of the GaN, we grew an $\text{Al}_{0.28}\text{Ga}_{0.78}\text{N}$ layer with a thickness of 25 nm on the GaN and demonstrated that the AlGaIn/GaN heterostructure with an area of two centimetres square has a two-dimensional electron gas mobility of $1,100 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ with a sheet carrier density of $1 \times 10^{13} \text{ cm}^{-2}$ at room temperature (293 K). Taken together, the AFM, X-ray diffraction, cross-sectional TEM, X-ray rocking curve, CBED and mobility confirm that single-crystal wurtzite GaN of device quality grows on the h-BN layers. In addition, a flat single-crystal GaN layer can be grown on a single-crystal AlGaIn layer on the h-BN (Supplementary Fig. 4). The single crystal h-BN layer works as the release layer. The AlN or AlGaIn is the buffer, and the thick GaN film ensures device quality in spite of the large lattice mismatch.

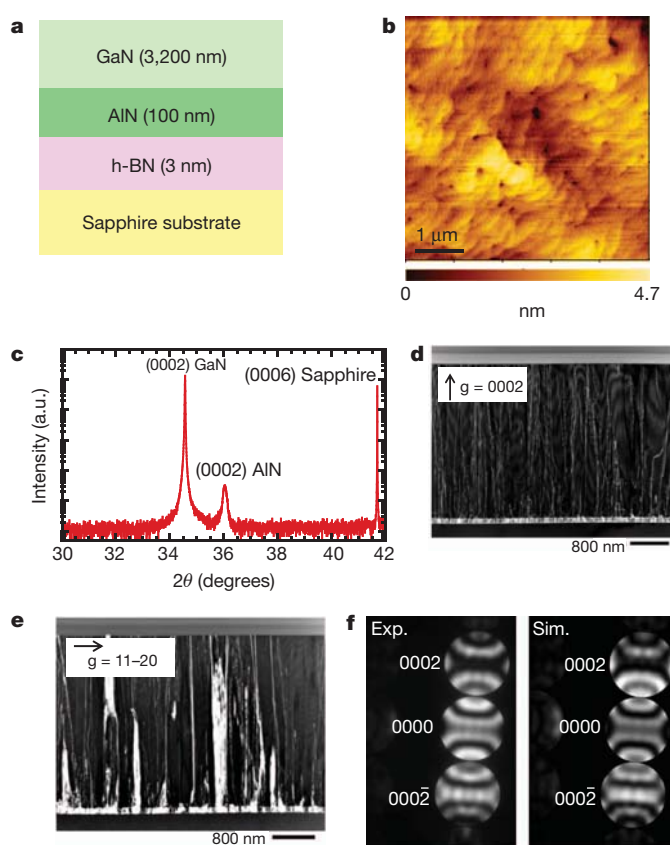


Figure 2 | Flat single-crystal wurtzite GaN is grown on an AlN layer on the h-BN layer. **a**, Schematic illustration. **b**, AFM image of the surface of the GaN film. **c**, X-ray diffraction using the $2\theta/\omega$ configuration for the GaN film. a.u., arbitrary units. **d**, **e**, Weak beam dark-field TEM images. **g** is the reciprocal lattice vector, and the arrows indicate the direction of the vectors. The numbers are Miller indices. **f**, Experimental and simulated CBED patterns. The numbers are Miller indices.

Before fabricating the LED structure, we grew an InGaIn/GaN MQW structure to investigate its optical properties. The photographs in Fig. 3a and b show the transferred AlGaIn/GaN heterostructure, approximately 2 cm square, and the MQW structure, approximately 5 mm square, on the indium sheets attached to the foreign sapphire substrates, respectively. We can see the surface of the indium sheet because the AlGaIn and the MQW are transparent. The size of the transferred area can be controlled by the adhesive sheet sizes. It takes only a few seconds to release the host sapphire substrate from the AlGaIn/GaN and MQW structure, which is an unusually speedy process compared to laser lift-off. The r.m.s. roughness of the indium sheet is 12 nm over an area of $5 \mu\text{m} \times 5 \mu\text{m}$. Some protrusions from the indium sheet are clearly visible, indicating that the AlGaIn/GaN and MQW structures are mechanically released from the native sapphire substrate, as illustrated in Fig. 1. No cracks were observed in the transparent AlGaIn/GaN structure up to maximum size of about 1 cm square, suggesting that the mechanical release process using the h-BN layer ensures minimal crack formation. The AFM image in Fig. 3c shows that the surface of the transferred structure is flat, with r.m.s. roughness of 0.95 nm over an area of $5 \mu\text{m} \times 5 \mu\text{m}$, which is considerably lower than that of surfaces separated by laser lift-off techniques^{25,26} (typically 10–60 nm). X-ray photoelectron spectroscopy spectra confirm that the separation actually occurs within the h-BN release layer along the plane of B and N bonding (Supplementary Fig. 5). X-ray diffraction (using the $2\theta/\omega$ configuration) measurements of the InGaIn/GaN MQW before and after the transfer showed satellite peaks from the MQW up to first order, along with those for GaN (0002), AlGaIn (0002) and sapphire (Fig. 3d). This indicates that the

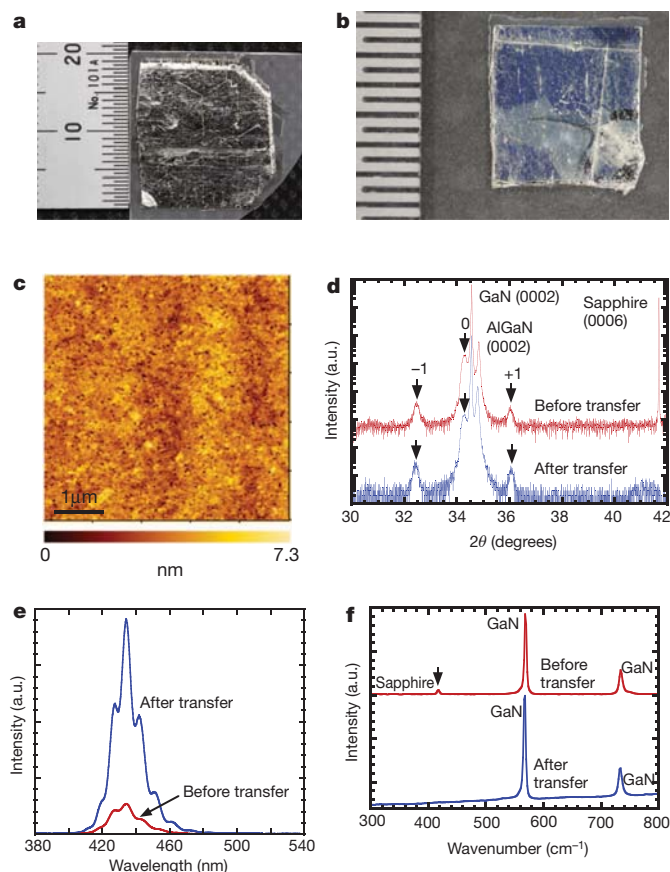


Figure 3 | Structural and optical properties of an InGaIn/GaN MQW structure before and after transfer. **a**, Photograph of the transferred AlGaIn/GaN structure attached to the indium sheet on the foreign sapphire substrate. The smallest division in the scale at left is 500 μm . **b**, Photograph of the transferred MQW structure attached to the indium sheet on the foreign sapphire substrate. One division in the scale at left is 500 μm . **c**, AFM image of the surface of the transferred structure. **d**, X-ray diffraction using the $2\theta/\omega$ configuration for the InGaIn/GaN MQW before and after the transfer. **e**, Photoluminescence spectra at room temperature for the MQW before and after transfer. Excitation wavelength, 375 nm. **f**, Raman spectra for the MQW at room temperature before and after transfer.

MQW period (the sum of well and barrier thickness) is 5.5 nm and that the In content in the InGaIn well is 15%. In Fig. 3d, the transferred MQW structure exhibits satellite peaks up to first order with almost the same intensities as those before the release.

To check for degradation of the nitride semiconductor that might be created by release and mounting, we performed room temperature photoluminescence measurements on the MQW before and after the transfer. The excitation source was an InGaIn-based semiconductor laser diode with an emission wavelength of 375 nm. Typical photoluminescence spectra from the MQW before and after transfer show strong luminescence at almost the same peak wavelength of 434 nm; however, the intensity after the transfer is comparable to or stronger than that before the transfer (Fig. 3e). We also grew the same InGaIn/GaN MQW structure on a GaN film using a standard low-temperature AlN buffer layer on sapphire substrate and performed room-temperature photoluminescence using a He-Cd laser. The luminescence intensity from the as-grown MQW structure on the h-BN release layer was comparable to or higher than that from the one grown on the low-temperature AlN buffer layer. The Raman spectrum for the MQW before the transfer displays the E_2 mode at 569 cm^{-1} and the A_1 mode from the GaN film, and shows the Raman mode from the host sapphire substrate as well (Fig. 3f). The Raman spectrum of the MQW after the transfer shows the E_2 mode from the GaN film at 567 cm^{-1} with a downshift of 2 cm^{-1} (Fig. 3f), indicating that the

GaN before and after the release is compressively strained and unstrained^{27,28}, respectively. The X-ray diffraction, photoluminescence and Raman experiments demonstrate that the MQW structure retains its original crystal quality, even after the transfer processes.

Next, we describe the electroluminescence emitted from the transferred LED (Supplementary Fig. 6a) at room temperature. For comparison, the same MQW LED structure was grown on a typical low-temperature AlN buffer layer on a sapphire substrate and the conventional MQW LED was fabricated without lift-off (Supplementary Fig. 6b). Current–voltage characteristics of the transferred LED show clear rectification (Fig. 4a). In the electroluminescence spectra of the transferred LED and conventional LED with currents ranging from 10 to 50 mA, the intensities of the electroluminescence increase with current (Fig. 4b and c) and the electroluminescence intensities from the transferred MQW LED were comparable to or higher than the intensities from the conventional MQW LED on the low-temperature AlN buffer layer at the same current. These higher electroluminescence intensities are caused by reflection from the back-side contact indium. The comparable intensities and almost the same FWHM of the electroluminescence of the conventional and transferred LED indicate that the MQW preserves its original quality after the transfer. We further succeeded in transferring a vertical-type LED (Supplementary Fig. 6c). This LED emits blue light at room temperature (Fig. 4d).

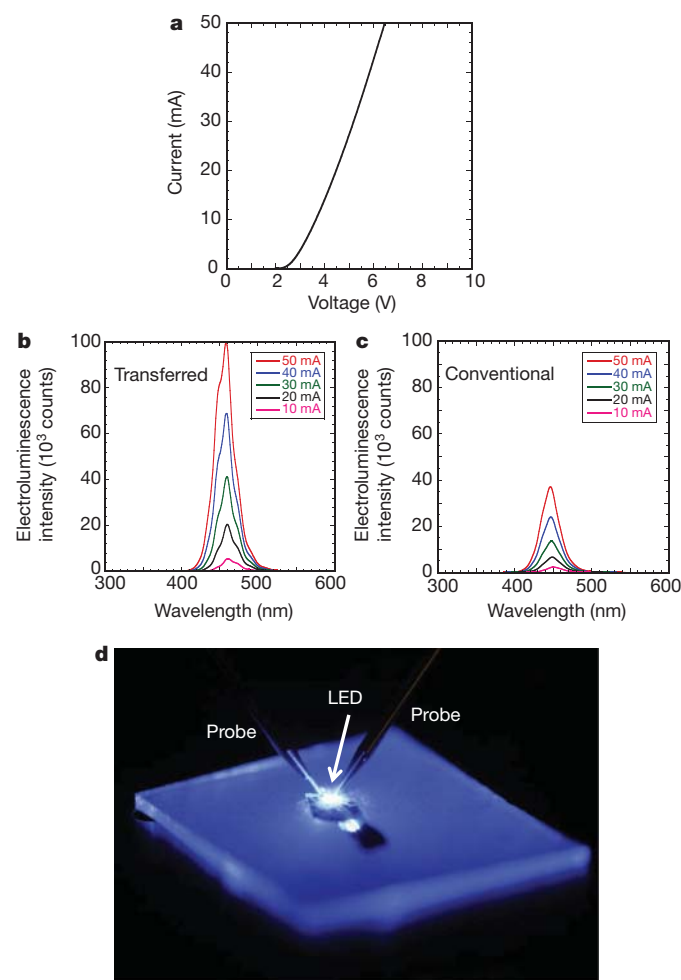


Figure 4 | Electroluminescence of transferred and conventional MQW LED at room temperature. **a**, Current–voltage characteristics of the transferred LED. **b**, Electroluminescence spectra from the transferred LED with currents ranging from 10 to 50 mA. **c**, Electroluminescence spectra from the conventional LED with currents ranging from 10 mA to 50 mA. **d**, Optical image of the blue-light electroluminescence from the transferred vertical-type LED.

Finally, we fabricated a battery-powered LED prototype. The released LED is about 2 mm square and 3.4- μm thick. A schematic cross-section of the LED prototype is shown in Fig. 5a. First, two T-shaped Pd/Au electrodes were deposited on two commercially available laminate films. Then, the released LED structure with the Pd/Au electrodes was mounted on the T-shaped electrode of one laminate film and sandwiched with the other T-shaped electrode of the other laminate film using the indium contact layer. The released LED was thereby sandwiched between the two laminates, which were sealed together by heat. Figure 5b shows a top-view photograph of the LED prototype, which is 18 mm long and 22 mm wide. The side-view of the LED prototype anchored with tweezers shows a thickness of about 200 μm (Fig. 5c), which can offer flexible and portable applications. The LED prototype emitted violet-blue electroluminescence at room temperature (Fig. 5d).

Compared with conventional laser lift-off, mechanical release using h-BN has several advantages. First, this approach requires no lift-off equipment and no chemical etchant, leading to considerable cost reduction. Second, the transfer in this approach is completed within several seconds because the separation harnesses the van der Waals forces of the h-BN release layer. In contrast, laser lift-off requires laser scanning with additional adjustment of the laser. Third, the process is simple: the mechanical force releases GaN-based structures in only one step. In laser lift-off, the laser must be scanned step-by-step, then the irradiated samples must be heated to the melting point of Ga, and finally the GaN sample after laser irradiation must be cleaned with a dilute acid (such as HCl) to remove material residues. The separated surface after the mechanical release is essentially flat, with an r.m.s.

roughness of 1 nm, which is considerably lower than that obtained using laser lift-off. The process we report here opens the way to releasing and transferring a wide range of nitride semiconductor devices to large-area, flexible and affordable substrates.

METHODS SUMMARY

The h-BN release layer, the AlGaIn/GaN heterostructure, the InGaIn/GaN MQW and the MQW LED structures were grown by MOVPE. Triethylboron, trimethylgallium, trimethylaluminum, triethylgallium and trimethylindium were the group III sources and ammonia was the group V source. Silane and bis-cyclopentadienyl-magnesium were the n-type and p-type dopant gases, respectively, with hydrogen or nitrogen carrier gas. Electron beam deposition provided Au, Pd and Al electrodes. A high-resolution X-ray diffractometer (Philips X'Pert System) with a copper target was used to evaluate the structural quality and perform X-ray diffraction pole figure measurement. The photoluminescence measurements were performed with an InGaIn-based laser at a wavelength of 375 nm and a He-Cd laser at room temperature. We used Raman scattering (Renishaw system) with an excitation laser wavelength of 534 nm. We used an electron cyclotron resonance plasma etching system with chlorine to fabricate LEDs.

Received 17 May 2011; accepted 14 February 2012.

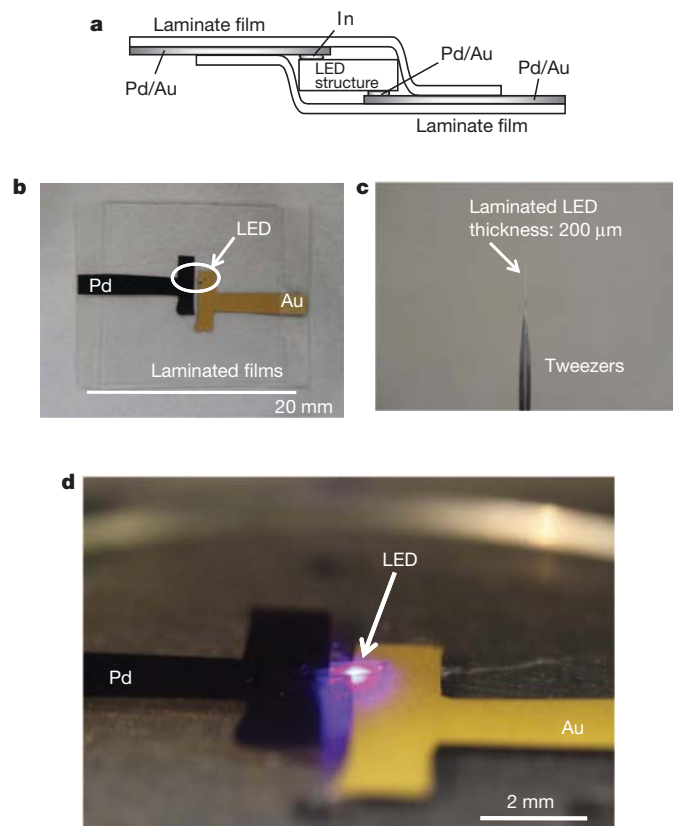


Figure 5 | Electroluminescence of LED prototype at room temperature. a, Schematic cross-section of the LED prototype. b, c, Top-view (b) and side-view (c) photographs of the LED prototype. The T-shaped black Pd electrode surface on one laminate film is on the left-hand side; the gold electrode on the other is on the right-hand side. The LED structure was mounted between the top of the two T-shaped electrodes. d, Optical image of the violet-blue electroluminescence from the LED prototype.

- Amano, H., Kito, M., Hiramatsu, K. & Akasaki, I. p-Type conduction in Mg-doped GaN treated with low-energy electron beam irradiation (LEEBI). *Jpn. J. Appl. Phys.* **28**, L2112–L2114 (1989).
- Nakamura, S. *et al.* InGaIn-based multi-quantum-well-structure laser diodes. *Jpn. J. Appl. Phys.* **35**, L74–L76 (1996).
- Asif Khan, M., Bhattarai, A., Kuznia, J. N. & Olson, D. T. High electron mobility transistor based on a GaN-Al_{0.1}Ga_{0.9}N heterojunction. *Appl. Phys. Lett.* **63**, 1214–1215 (1993).
- Wong, W. S., Sand, T. & Cheung, N. W. Damage-free separation of GaN thin films from sapphire substrates. *Appl. Phys. Lett.* **72**, 599–601 (1998).
- Kelly, M. K. *et al.* Large free-standing GaN substrates by hydride vapor phase epitaxy and laser-induced liftoff. *Jpn. J. Appl. Phys.* **38**, L217–L219 (1999).
- Bour, D. P. *et al.* Polycrystalline nitride semiconductor light-emitting diodes fabricated on quartz substrates. *Appl. Phys. Lett.* **76**, 2182–2184 (2000).
- Amano, H., Sawaki, N., Akasaki, I. & Toyoda, Y. Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer. *Appl. Phys. Lett.* **48**, 353–355 (1986).
- Nakamura, S. GaN growth using GaN buffer layer. *Jpn. J. Appl. Phys.* **30**, L1705–L1707 (1991).
- Kumakura, K. & Makimoto, T. Growth of GaN on sapphire substrates using novel buffer layers of ECR-plasma-sputtered Al₂O₃/graded-AlON/AlN/Al₂O₃. *J. Cryst. Growth* **292**, 155–158 (2006).
- Ahn, J.-H. *et al.* Heterogeneous three-dimensional electronics by use of printed semiconductor nanomaterials. *Science* **314**, 1754–1757 (2006).
- Rogers, D. J. *et al.* Use of ZnO thin films as sacrificial templates for metal organic vapor phase epitaxy and chemical lift-off of GaN. *Appl. Phys. Lett.* **91**, 071120 (2007).
- Lee, S. W. *et al.* Lattice strain in bulk GaN epilayers grown on CrN/sapphire template. *Appl. Phys. Lett.* **94**, 082105 (2009).
- Tripathy, S. *et al.* InGaIn/GaN light emitting diodes on nanoscale silicon on insulator. *Appl. Phys. Lett.* **91**, 231109 (2007).
- Lee, K. J. *et al.* Bendable GaN high electron mobility transistors on plastic substrates. *J. Appl. Phys.* **100**, 124507 (2006).
- Tamboli, A. C. *et al.* Room-temperature continuous-wave lasing in GaN/InGaIn microdisks. *Nature Photon.* **1**, 61–64 (2007).
- Cho, C.-Y. *et al.* Growth and separation of high quality GaN epilayer from sapphire substrate by lateral epitaxial overgrowth and wet chemical etching. *Appl. Phys. Express* **4**, 012104 (2011).
- Zang, K. Y. *et al.* A new method for lift-off of III-nitride semiconductors for heterogeneous integration. *Nanoscale Res. Lett.* **5**, 1051–1056 (2010).
- Chung, K., Lee, C.-H. & Yi, G.-C. Transferable GaN layers grown on ZnO-coated graphene layers for optoelectronic devices. *Science* **330**, 655–657 (2010).
- Novoselov, K. S. *et al.* Electric field effect in atomically thin carbon films. *Science* **306**, 666–669 (2004).
- Boo, J.-H., Rohr, C. & Ho, W. MOCVD of BN and GaN thin films on silicon: new attempt of GaN growth with BN buffer layer. *J. Cryst. Growth* **189–190**, 439–444 (1998).
- Kobayashi, Y. & Akasaka, T. Hexagonal BN epitaxial growth on (0001) sapphire substrate by MOVPE. *J. Cryst. Growth* **310**, 5044–5047 (2008).
- Kobayashi, Y., Tsai, C.-L. & Akasaka, T. Optical band gap of h-BN epitaxial film grown on c-plane sapphire substrate. *Phys. Status Solidi C* **7**, 1906–1908 (2010).
- Kobayashi, K. *et al.* X-ray rocking curve determination of twist and tilt angles in GaN films grown by an epitaxial-lateral-overgrowth technique. *Jpn. J. Appl. Phys.* **38**, L611–L613 (1999).
- Metzger, T. *et al.* Defect structure of epitaxial GaN films determined by transmission electron microscopy and triple-axis X-ray diffractometry. *Phil. Mag. A* **77**, 1013–1025 (1998).
- Ueda, T., Ishida, M. & Yuri, M. Laser lift-off of very thin AlGaIn film from sapphire using selective decomposition of GaN interlayer. *Appl. Surf. Sci.* **216**, 512–518 (2003).

26. Chu, C.-F. *et al.* Study of GaN light-emitting diodes fabricated by laser lift-off technique. *J. Appl. Phys.* **95**, 3916–3922 (2004).
27. Kisielowski, C. *et al.* Strain-related phenomena in GaN thin films. *Phys. Rev. B* **54**, 17745–17753 (1996).
28. Davydov, V. Y. *et al.* Phonon dispersion and Raman scattering in hexagonal GaN and AlN. *Phys. Rev. B* **58**, 12899–12907 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Oudah for help with AFM, X-ray diffraction and Raman measurements, T. Kitada for help with TEM observations, H. Ando for help with X-ray photoelectron spectroscopy, H. Gotoh for help with photoluminescence using the InGaN-based laser, Y. Yamauchi for help with LED processing, Y. Tanaiyasu for discussions about LED characterization, and Y. Krockenberger for reading and

technical support of the manuscript. We also thank H. Yamamoto and I. Yokohama for their encouragement throughout this study.

Author Contributions Y.K. designed the experiments, grew the MQW and LED structures and characterized their structural and optical properties. K.K. fabricated the electrode, released and transferred the LED, and performed the current–voltage and electroluminescence measurements. T.A. also grew the MQW and LED structures. T.M. managed the whole experimental process. Y.K. and T.M. proposed this research and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.K. (kobayashi.yasuyuki@lab.ntt.co.jp).

Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability

Ben B. Booth¹, Nick J. Dunstone^{1*}, Paul R. Halloran^{1*}, Timothy Andrews¹ & Nicolas Bellouin¹

Systematic climate shifts have been linked to multidecadal variability in observed sea surface temperatures in the North Atlantic Ocean¹. These links are extensive, influencing a range of climate processes such as hurricane activity² and African Sahel^{3–5} and Amazonian⁵ droughts. The variability is distinct from historical global-mean temperature changes and is commonly attributed to natural ocean oscillations^{6–10}. A number of studies have provided evidence that aerosols can influence long-term changes in sea surface temperatures^{11,12}, but climate models have so far failed to reproduce these interactions^{6,9} and the role of aerosols in decadal variability remains unclear. Here we use a state-of-the-art Earth system climate model to show that aerosol emissions and periods of volcanic activity explain 76 per cent of the simulated multidecadal variance in detrended 1860–2005 North Atlantic sea surface temperatures. After 1950, simulated variability is within observational estimates; our estimates for 1910–1940 capture twice the warming of previous generation models but do not explain the entire observed trend. Other processes, such as ocean circulation, may also have contributed to variability in the early twentieth century. Mechanistically, we find that inclusion of aerosol–cloud microphysical effects, which were included in few previous multimodel ensembles, dominates the magnitude (80 per cent) and the spatial pattern of the total surface aerosol forcing in the North Atlantic. Our findings suggest that anthropogenic aerosol emissions influenced a range of societally important historical climate events such as peaks in hurricane activity and Sahel drought. Decadal-scale model predictions of regional Atlantic climate will probably be improved by incorporating aerosol–cloud microphysical interactions and estimates of future concentrations of aerosols, emissions of which are directly addressable by policy actions.

An understanding of North Atlantic sea surface temperature (NASST) variability is critical to society because historical Atlantic temperature changes are strongly linked to the climate, and its impacts, in neighbouring continental regions. For example, strong links between NASST variability and periods of African Sahel drought are found in observations^{4,13} and physical climate models^{3,5,14}. Similar covariation between NASSTs and rainfall in eastern South America has been found⁵, as have links to changes in both mean rainfall¹⁵ and rainfall extremes¹⁶, Atlantic hurricane activity^{2,10,14} and European summer climate⁸. These changes are not solely limited to the regions bordering the Atlantic, but also have links to Indian monsoon rainfall¹⁴, Arctic and Antarctic temperatures¹⁷, Hadley circulation¹, El Niño/Southern Oscillation¹⁸ and relationships between El Niño/Southern Oscillation and the Asian monsoon¹⁹.

A link between multidecadal variability in NASST and circulation changes internal to the ocean was first proposed in 1964 (ref. 20) and later named the Atlantic Multidecadal Oscillation²¹. This variability is often characterized as the detrended NASST between the equator and latitude 60° N (longitude 7.5–75° W; ref. 8). Although it has recently been questioned²², the present consensus remains that most of the observed Atlantic temperature variations occur in response to the

ocean's internal variability. This picture emerged from general circulation models, a number of which inherently produce multidecadal Atlantic variability in the absence of external climate forcing⁷ and, when considered together as a multimodel mean, have shown little evidence of forced changes projecting onto the NASSTs^{6,9}. Observationally, this interpretation has been accepted because the Atlantic temperature changes seem to be oscillatory, both around any secular long-term trend and when calculated as anomalies from the global-mean change.

Motivated by the recent identification of the importance of aerosol process complexity in interhemispheric Atlantic temperature changes²³, apparent aerosol correlation^{1,11} and volcanic modulation of Atlantic variability²², we use new general circulation model simulations to question whether the CMIP3 (Climate Model Intercomparison Project phase 3) models contained the complexity necessary to represent a forced Atlantic Multidecadal Oscillation^{7,9}. We use HadGEM2-ES (the Hadley Centre Global Environmental Model version 2 Earth System configuration²⁴), a next-generation CMIP5 (Climate Model Intercomparison Project phase 5) model, which represents a wider range of Earth system processes (in particular aerosol interactions²⁵) than do CMIP3 models.

To separate internal variability from forced changes, we present climate model ensemble-mean NASSTs, averaged over parallel model simulations started from different initial conditions²⁶. If external forcing dominates the NASST evolution then ensemble members will evolve in phase and thus combine to produce a robust ensemble-mean response. If internal ocean dynamics dominate then each member will evolve separately and the resulting ensemble mean will show little residual variation around the underlying warming trend. This approach allows identification of physical mechanisms linking forced changes to Atlantic temperatures and was used in previous CMIP3 studies^{6,9}.

In Fig. 1a, we reproduce the multimodel-mean NASST response of the six CMIP3 models used in ref. 9 (ENS1, blue) and the eleven models used in ref. 6 (ENS2, green) (Supplementary Table 2). The observations (Fig. 1) show marked multidecadal variations. The multimodel-mean responses in both ENS1 and ENS2 do capture the underlying trend through the century; they capture only weak multidecadal variability. For example, the ensembles' 1950–1975 cooling is only a small fraction of the observed value (Fig. 1a and Supplementary Fig. 4). Therefore, the unexplained multidecadal signal was previously attributed to internal ocean variability^{6,9}.

By contrast, HadGEM2-ES (Fig. 1b) reproduces much more of the observed NASST variability (correlation, 0.65; 75% of detrended standard deviation (smoothed over 10-yr intervals to highlight multidecadal component)). The post-1950s cooling and subsequent warming now falls within the observed trends (Supplementary Table 1). Observed warming in the earlier period (1910–1940) is larger than simulated by HadGEM2-ES (Fig. 1b and Supplementary Table 1); however, these new simulations capture roughly twice the early-twentieth-century warming of previous CMIP3 generation models.

¹Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.

*These authors contributed equally to this work.

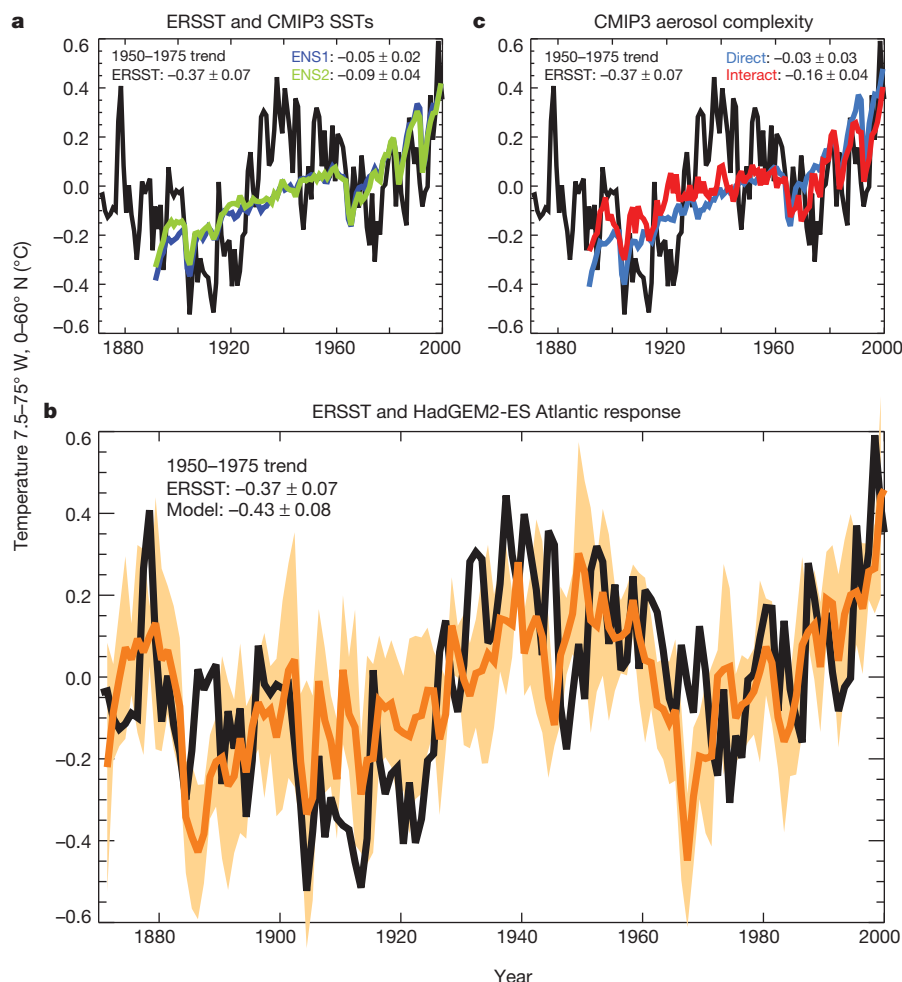


Figure 1 | Atlantic surface temperatures. Comparison of the area-averaged North Atlantic SSTs (defined as 7.5–75° W and 0–60° N), relative to the 1901–1999 average, of an observational estimate (the US National Oceanic and Atmospheric Administration's Extended Reconstructed SST²⁷ (ERSST), black) and two published^{6,9} CMIP3 model composites (ENS1, blue; ENS2, green; **a**); the HadGEM2-ES model (orange; shading represents 1 s.d. of the model ensemble spread; **b**); and two recomposites from CMIP3, the first with models

that represent only direct aerosol (mean of five contributing models, red) and the second with models representing both indirect effects interactively (three models, blue) (**c**). In all panels, trends between 1950–1975 (K per decade) are shown. The error estimates are based on the s.d. of the 25 trends between a 5-yr period (1948–1952) at the start of this interval and a 5-yr period (1973–1977) at the end. All data have been latitude-weighted when calculating area averages.

This points to a larger forced role in this period. Other processes not represented by our ensemble-mean response (such as ocean dynamical changes) may also contribute to this early trend.

In examining why the HadGEM2-ES ensemble reproduces the observed NASST variability better than previous multimodel studies have done^{6,9} (Fig. 1a, b), we can discount the possibility that the HadGEM2-ES variability is predetermined, because the initial conditions were selected to sample different phases of Atlantic variability²⁶. Furthermore, an additional HadGEM2-ES ensemble that omits changes in aerosol emissions neither has the same multidecadal variability as the all-forcings ensemble nor reproduces the observed NASSTs (Fig. 2a).

Replication of a large fraction of the observed NASST variability by HadGEM2-ES allows us to identify forcings and mechanisms, consistent with the observed variability, within the model framework. Variability of ensemble-mean NASST from historical simulations including time-varying aerosol emissions is strongly correlated with variability in simulated net surface shortwave radiation (Fig. 2b), which in turn has the same temporal structure as variability in aerosol optical depth changes (Fig. 2c) and periods of volcanic activity (Fig. 2d). Other terms in the surface heat budget (Supplementary Fig. 2) have a role in the simulated NASST change. However, it is the surface shortwave component that produces the dominant multidecadal variations.

Volcanoes and aerosols respectively explain 23 and 66% of the temporal (10-yr-smoothed) multidecadal variability of the detrended NASST (Supplementary Fig. 5). Combining both contributions explains 76% (80% after inclusion of mineral dust aerosols) of the simulated variance. Inclusion of mineral dust processes may potentially be important because emissions are known to respond to North-Atlantic-driven changes in Sahel rainfall, and thus represent an important positive feedback on NASSTs in the real world¹². The lack of a multidecadal dust signal (Supplementary Information) in HadGEM2-ES simulations suggests that we are likely to be underestimating the magnitude of the forced Atlantic response.

The volcanic influence on Atlantic variability has been demonstrated previously^{12,22}. We focus on the anthropogenic aerosol component of the shortwave changes identified here as driving the model's multidecadal NASST variability. Aerosol concentration changes influence the spatial response (Fig. 3) of NASST as well as its temporal evolution. Prevailing winds advect aerosols emitted in industrial North America in a band across the North Atlantic that mixes with polluted air masses over Europe before being transported by trade winds south and west. The large-scale pattern of shortwave change is explained by the effect of cloud microphysical response to these changes in aerosol concentration. The shortwave variability largely occurs where aerosol changes coincide with large-scale cloud distribution. On a regional

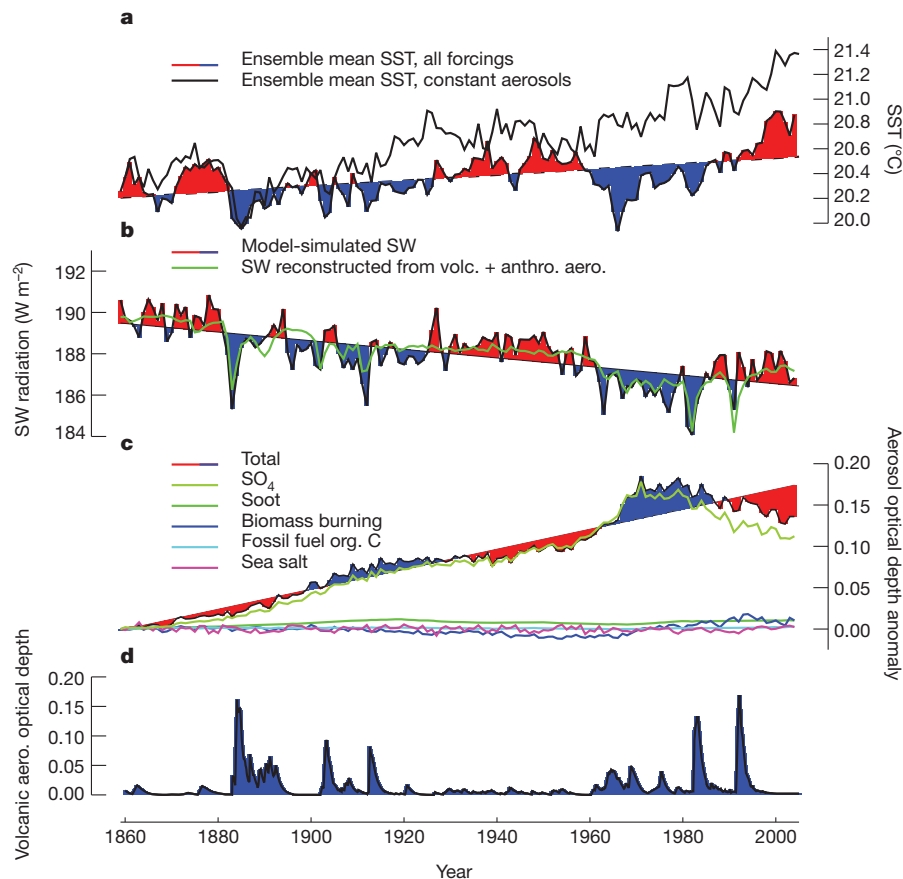


Figure 2 | External forcing of surface temperature and surface shortwave radiation linked to aerosol and volcanic changes. **a**, Ensemble-mean NASST (7.5–75° W and 0–60° N) simulated by the HadGEM2-ES model considering all available external climate forcings (red/blue), and all forcings except anthropogenic aerosol emissions (black). **b**, Ensemble-mean shortwave (SW) radiation entering ocean (red/blue) alongside reconstruction of this shortwave radiation based on a linear relationship with total anthropogenic aerosol and

volcanic optical depth changes (green). The linear model explains 79% of the simulated variance. **c**, Change in total (red/blue shading) and individual species (coloured lines) of anthropogenic aerosol optical depth (degree of absorption/scattering) over the North Atlantic. **d**, Volcanic optical depth from ref. 30 as implemented in HadGEM2-ES simulations. In **a–c**, red and blue shading represents values above or below (or vice versa) a least-squares linear fit to the data.

scale, coupled processes, such as temperature feedbacks, can lead to an enhanced local response (Supplementary Information). The same horseshoe-shaped signature is seen in shortwave and NASST variability (Fig. 3). The map of NASST change between warm and cold phases of multidecadal variability is consistent with the observed variations in SSTs (Fig. 3).

So far consistent spatial and temporal changes between aerosol burden, shortwave and NASST have been presented. It is not clear, however, whether these changes are externally forced by aerosols or are mediated by ocean circulation. Here we present a parallel simulation of the historical period, driven by identical emission and concentration changes, but with the SSTs explicitly fixed at their 1860 climatological values. The shortwave changes arising in this parallel experiment share the temporal structure and magnitude of the shortwave changes from our standard historical simulations (Fig. 4a). By explicitly removing any feedback from SST change on shortwave, we demonstrate that simulated historical shortwave variability arises directly from aerosol and volcanic forcing of the surface radiation and is not mediated by ocean circulation change.

One of the reasons why the role of aerosols in driving multidecadal variability has not previously been identified is the level of aerosol physics represented in climate models at the time of the CMIP3 multi-model comparison project (Supplementary Table 2). Although all the CMIP3 models represented the direct effect of aerosols on shortwave radiation, most omitted or only partly represented the indirect aerosol effects²³.

Recently, albedo differences in CMIP3 aerosol representation have been shown to be important for simulating SST changes²³: models that represent indirect aerosol effects capture more of the observed Atlantic interhemispheric change than those that do not. Recompositing the models used in Fig. 1a into those with only direct aerosols effects, and those that also include the first indirect effect interactively, shows some evidence of multidecadal variability (Fig. 1c), illustrating that aerosol–cloud microphysical processes have a role even in previous-generation models. These models do not, however, reproduce the magnitude of the multidecadal NASST of HadGEM2-ES.

In HadGEM2-ES, the aerosol indirect effects account for 80% or more of the total aerosol forcing in the North Atlantic region (Fig. 4b and Supplementary Fig. 1). Although there is some discussion of the magnitude of the indirect effects^{28,29}, omission of these processes will lead to an underestimation of the modelled aerosol impact on the NASST. Looking at the relative roles of the first and second indirect effects (using changes in optical depth and cloud effective radius as respective metrics for these effects; Fig. 4c), we see a more pronounced response to early-twentieth-century variations for the indirect effect (effective radius) due to higher sensitivity of cloud albedo changes to changes in aerosol number in cleaner conditions³. In all, the inclusion of aerosol indirect effects in HadGEM2-ES magnifies the shortwave and, hence, the NASST response to aerosols, as well as influencing the spatial and temporal character of the historical multidecadal variability. We also note that although some climate models (such as HadGEM2-ES) reproduce the observed sensitivity of cloud albedo to changes in

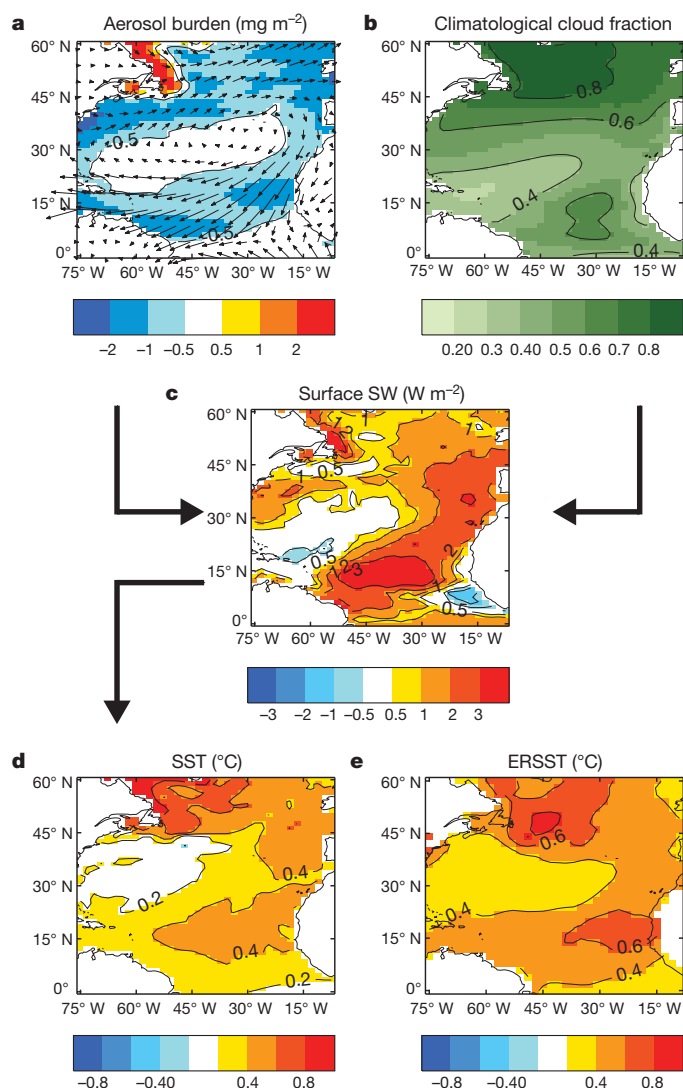


Figure 3 | Differences in spatial response between warm and cold Atlantic phases. Differences between the average of warm years (warmest third of the data) and the average of cold years (coldest third of the data), after the data has been detrended, show how patterns of multidecadal variability in aerosol burden (a) interact with the climatological cloud field (b) to influence the pattern of net surface shortwave radiation change (c) and, hence, NASSTs (d). The pattern of aerosol burden changes is linked to emissions in industrial North America and Europe by the climatological wind field (the direction and magnitude of which is indicated by the arrows). The exception is the localized increase in aerosol burden (Canadian coast) driven by increases in sea salt aerosols (warm years reduce sea ice extent in this area). The warm phase/cold phase SST pattern simulated by the model (d) agrees well with the observed change (e). Vertical axes show latitude; horizontal axes show longitude.

aerosol optical depth in maritime regions, not all parameterizations of aerosol indirect effects do so (Fig. 2e in ref. 29).

We have shown that volcanic and aerosol processes can drive pronounced multidecadal variability in historical NASST, which leads to improved (for the early twentieth century) or reproduces (for the later period) the observed historical trends. In these simulations, it is the inclusion of aerosol indirect effects that allows us to capture the magnitude and the temporal and spatial structure of SST variability. Our results show that volcanoes and, crucially (from a policy and climate impact perspective), anthropogenic emissions of aerosols can drive NASST variability resembling that which is observed. This work suggests that we need to reassess the current attribution to natural ocean variability of a number of prominent past climate impacts linked to NASSTs, such as Sahel drought.

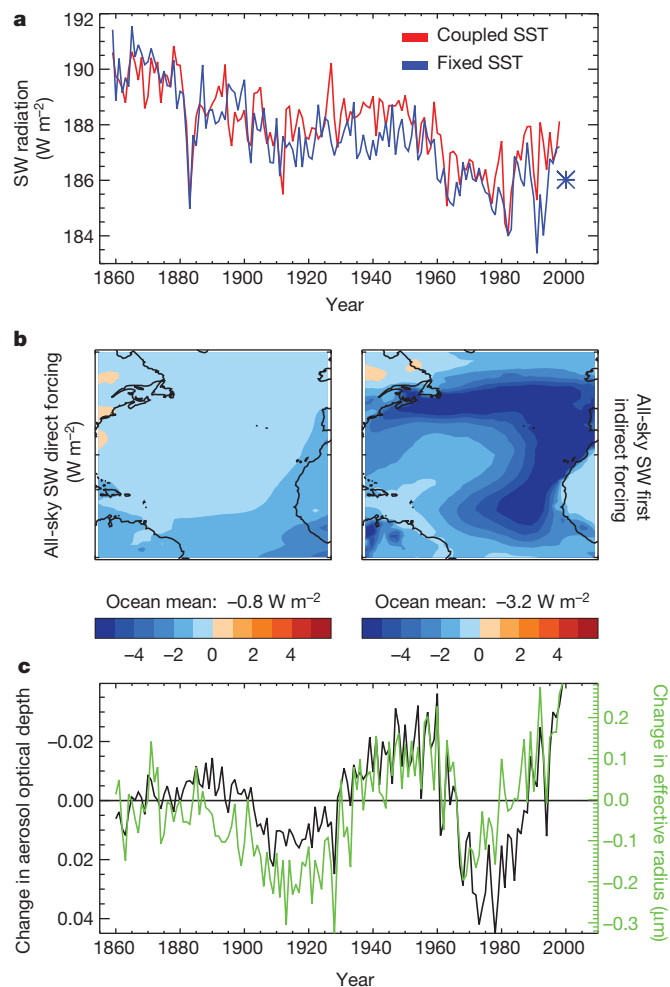


Figure 4 | Magnitude and origin of forced changes in net surface shortwave radiation. a, Ensemble-mean time evolution of surface shortwave in the North Atlantic region (red; see also Fig. 2b) and the surface shortwave from a parallel simulation in which the SSTs were held fixed at climatological values from 1860 (blue). The comparison shows that these shortwave changes are externally forced. These long-term trends in surface shortwave match the changes observed in a snapshot experiment (2000 aerosol emissions, 1860 SSTs; blue asterisk), indicating that the underlying trend is consistent with that expected from aerosol changes alone. b, Spatial patterns of surface shortwave forcing arising from the 2000 direct effect (left) and the first indirect effect (right) of aerosols (Supplementary Information), which illustrate the dominant role of indirect effects in the total forcing and the spatial distribution. c, Detrended time series of aerosol optical depth (black) and cloud-top effective radius (green) from the coupled simulations, which are indicators of the temporal evolution of direct and, respectively, indirect effects. Although the variations in effective radius are largely in phase with those in optical depth, there is greater divergence (implying a larger role for indirect effects) in the early historical period.

METHODS SUMMARY

The climate model used in this study is HadGEM2-ES²⁴. This model is notable for the number of climate–biogeochemical interactions that are interactively calculated rather than specified in advance. Of relevance to this work, HadGEM2-ES models the supply of oxidants, an important component for aerosol formation, and mineral dust aerosols interactively, and yields improved predictions of biomass and carboniferous aerosol properties. Source terms for natural aerosols (or precursors) and mineral dust are also modelled interactively. Each simulation in the ensemble is forced with driving data (greenhouse gases, aerosols, volcanoes, land use and solar changes) based on historical data sets compiled for CMIP5 simulations. These data sets and their implementation within this model are extensively documented in ref. 26. Volcanic forcing is prescribed in latitudinal bands. Over the North Atlantic, the magnitude of optical depth changes are prescribed individually for the bands spanning 0–30° N and 30–90° N, capturing the differences between tropical and extratropical volcanoes. Individual members of

the ensemble were initiated from a control simulation using start points located 50 yr apart²⁶.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 September 2011; accepted 8 February 2011.

Published online 4 April 2012.

- Baines, P. & Folland, C. K. Evidence for a rapid global climate shift across the late 1960s. *J. Clim.* **20**, 2721–2744 (2007).
- Goldenberg, S. B., Landsea, C. W., Nunez, A. M. M. & Gray, W. M. The recent increase in Atlantic hurricane activity: causes and implications. *Science* **293**, 474–479 (2001).
- Ackerley, D. *et al.* Sensitivity of twentieth-century Sahel rainfall to sulfate aerosol and CO₂ forcing. *J. Clim.* **24**, 4999–5014 (2011).
- Hoerling, M. J., Hurrell, J., Eischeid, J. & Phillips, A. Detection and attribution of 20th century Northern and Southern African rainfall change. *J. Clim.* **19**, 3989–4008 (2006).
- Knight, J. R., Folland, C. K. & Scaife, A. A. Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.* **33**, L17706 (2006).
- Knight, J. R. The Atlantic Multidecadal Oscillation inferred from the forced climate response in coupled general circulation models. *J. Clim.* **22**, 1610–1625 (2009).
- Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M. & Mann, M. E. A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Res. Lett.* **32**, L20708 (2005).
- Sutton, R. T. & Hodson, D. L. R. Atlantic Ocean forcing of North American and European summer climate. *Science* **309**, 115–118 (2005).
- Ting, M., Kushnir, Y., Seager, R. & Li, C. Forced and internal twentieth-century SST trends in the North Atlantic. *J. Clim.* **22**, 1469–1481 (2009).
- Trenberth, K. E. & Shea, D. J. Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.* **33**, L12704 (2006).
- Mann, M. E. & Emanuel, K. A. Atlantic hurricane trends linked to climate change. *Eos* **87**, 233–244 (2006).
- Evan, A. T., Vimont, D. J., Heidinger, A. K., Kossin, J. P. & Bennartz, R. The role of aerosols in the evolution of tropical North Atlantic Ocean temperature anomalies. *Science* **324**, 778–781 (2009).
- Folland, C. K., Palmer, T. N. & Parker, D. E. Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature* **320**, 602–607 (1986).
- Zhang, R. & Delworth, T. L. Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.* **33**, L17712 (2006).
- Enfield, D. B., Nunez, A. M. M. & Trimble, J. P. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophys. Res. Lett.* **28**, 2077–2080 (2001).
- Curtis, S. The Atlantic multidecadal oscillation and extreme daily precipitation over the US and Mexico during the hurricane season. *Clim. Dyn.* **30**, 343–351 (2008).
- Chylek, P., Folland, C. K., Lesins, G. & Dubey, M. K. Twentieth century bipolar seesaw of the Arctic and Antarctic surface air temperatures. *Geophys. Res. Lett.* **37**, L08703 (2010).
- Dong, B. W., Sutton, R. T. & Scaife, A. A. Multidecadal modulation of El Niño–Southern Oscillation (ENSO) variance by Atlantic Ocean sea surface temperatures. *Geophys. Res. Lett.* **33**, L08705 (2006).
- Chen, W., Dong, B. & Lu, R. Impact of the Atlantic Ocean on the multidecadal fluctuation of El Niño–Southern Oscillation–South Asian monsoon relationship in a coupled general circulation model. *J. Geophys. Res.* **115**, D17109 (2010).
- Bjerknes, J. Atlantic air-sea interaction. *Adv. Geophys.* **10**, 1–82 (1964).
- Kerr, R. A. A North Atlantic climate pacemaker for the centuries. *Science* **288**, 1984–1985 (2000).
- Otterå, O. H., Bentsen, M., Drange, H. & Suo, L. External forcing as a metronome for Atlantic multidecadal variability. *Nat. Geosci.* **3**, 688–694 (2010).
- Chang, C. Y., Chiang, J. C. H. & Wehner, M. F. Friedman, A. & Ruedy, R. Sulfate aerosol control of tropical Atlantic climate over the 20th century. *J. Clim.* **24**, 2540–2555 (2011).
- Collins, W. J. *et al.* Development and evaluation of an Earth-system model - HadGEM2. *Geosci. Model Dev.* **4**, 1051–1075 (2011).
- Bellouin, N. *et al.* Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate. *J. Geophys. Res.* **116**, D20206 (2011).
- Jones, C. D. *et al.* The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geosci. Model Dev.* **4**, 543–570 (2011).
- Smith, T. M., Reynolds, R. W., Peterson, T. C. & Lawrimore, J. Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). *J. Clim.* **21**, 2283–2296 (2008).
- Lohmann, U. & Feichter, J. Global indirect aerosol effects: a review. *Atmos. Chem. Phys.* **5**, 715–737 (2005).
- Quaas, J. *et al.* Aerosol indirect effects - general circulation model intercomparison and evaluation with satellite data. *Atmos. Chem. Phys.* **9**, 8697–8717 (2009).
- Sato, M., Hansen, J. E., McCormick, M. P. & Pollack, J. B. Stratospheric aerosol optical depths, 1850–1990. *J. Geophys. Res.* **98**, 22987–22994 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful for discussion and input from D. Smith, D. Sexton, J. Murphy, M. Palmer, C. Roberts and J. Knight during the analysis and writing of this paper. We acknowledge the modelling groups, the Program for Climate Model Diagnosis and Intercomparison and the WCRP's Working Group on Coupled Modelling for their roles in making available the WCRP CMIP3 multimodel data set. Support of this data set is provided by the Office of Science, US Department of Energy. The authors were supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101) and the EU FP7 THOR project.

Author Contributions B.B.B.B., N.J.D. and P.R.H. jointly led the analysis, produced figures and wrote the paper. T.A. and N.B. contributed additional experiments to diagnose different aspects of the aerosol forcing, provided technical and scientific insight and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.B.B.B. (ben.booth@metoffice.gov.uk).

METHODS

The central finding of this paper (that changes in volcanic and aerosol forcing are capable of driving variability in NASSTs much like that observed) is based on an ensemble of four HadGEM2-ES historical simulations (Methods Summary). The following parallel simulations were also used.

- An ensemble of three HadGEM2-ES simulations (parallel to the first three members of the all-forcings ensemble) with no changes in aerosol emissions. Results of this experiment are shown in Fig. 2 and Supplementary Fig. 2. These simulations used identical driving data to the standard historical ensemble, prescribing changes in emissions and concentrations based on the CMIP5 historical data sets. The exception is the anthropogenic aerosol emissions, which (along with the surface chemistry and consequent contribution to aerosol oxidation) were kept constant at their 1860 values. This ensemble provides a comparison with historical NASSTs where the historical changes in aerosol emissions did not take place.
- A parallel historical simulation of HadGEM2-ES in which the annual cycle of global SSTs are held at 1860 values, as calculated from the pre-industrial control simulation. Results of this experiment are shown in Fig. 4. This simulation used identical driving data to the standard historical ensemble and is designed to characterize the evolving nature of the historical forcings. Radiative forcing is the impact on the radiative balance resulting from a change or changes in the atmospheric constituents, or other external change (such as solar), before any SST-driven feedback on that change. This parallel run with fixed SSTs provides information on the shortwave changes (or the changes in any other radiative component) due directly to changes in atmospheric concentrations, explicitly removing any contribution arising from SST-driven feedbacks.

- As a companion experiment to the fixed-SST historical simulation (described above), snapshot experiments were carried out to assess the impact of aerosol changes alone between the beginning and end of this simulation. This set-up used SSTs, and all forcings other than anthropogenic aerosol emissions, held at their 1860 values. Anthropogenic aerosol emissions were set to either their 2000 or 1860 values. The comparison of the two allows an estimate to be made of the radiative changes (including shortwave) between 1860 and 2000, rather than, for example, as a feedback to the warming SSTs. The result of this experiment is shown in Fig. 4 alongside the fixed-SST historical simulation (in which other atmospheric consistent also varied).
- Three snapshot experiments in which the model code calculating the instantaneous radiation balance was run twice at each forward step of the model—once with the relevant aerosol process included and one without—to quantify separately the radiative impact of the direct and the first indirect effects of aerosols. The three snapshot experiments used aerosol emissions from 1950, 1980 and 2000 and calculated the surface shortwave radiative impact (forcing) of aerosols in those three years. These are presented in Supplementary Fig. 1 and the 2000 snapshot is included in Fig. 4. The value of these runs is that they allow the radiative impacts of these two aerosol effects to be compared.

The HadGEM2-ES model, like others before it⁹, captures SST variability in simulations unforced by external factors (fig. 20 in ref. 25), which, in unforced pre-industrial simulations, are strongly correlated with variability in Atlantic meridional overturning circulation (0.65, using 10-yr smoothing). However, variations in circulation are less important in the ensemble-mean variability of the historical simulations, where shortwave forcing dominates. This is discussed further in the section on shortwave changes and the surface heat budget in Supplementary Information.

Adaptation at the output of the chemotaxis signalling pathway

Junhua Yuan¹, Richard W. Branch¹, Basarab G. Hosu¹ & Howard C. Berg¹

In the bacterial chemotaxis network, receptor clusters process input^{1–3}, and flagellar motors generate output⁴. Receptor and motor complexes are coupled by the diffusible protein CheY-P. Receptor output (the steady-state concentration of CheY-P) varies from cell to cell⁵. However, the motor is ultrasensitive, with a narrow operating range of CheY-P concentrations⁶. How the match between receptor output and motor input might be optimized is unclear. Here we show that the motor can shift its operating range by changing its composition. The number of FliM subunits in the C-ring increases in response to a decrement in the concentration of CheY-P, increasing motor sensitivity. This shift in sensitivity explains the slow partial adaptation observed in mutants that lack the receptor methyltransferase and methylesterase^{7,8} and why motors show signal-dependent FliM turnover⁹. Adaptive remodelling is likely to be a common feature in the operation of many molecular machines.

The chemotaxis signalling pathway allows bacterial cells to sense and respond to changes in concentrations of chemical attractants or repellents^{1,2}. Binding of chemicals by receptors modulates the activity of an associated histidine kinase, CheA, thereby changing the level of phosphorylation of the response regulator, CheY. CheY-P binds to FliM, a component of the switch complex at the base of the flagellar motor and modulates the direction of motor rotation. A phosphatase, CheZ, dephosphorylates CheY-P. The chemotaxis pathway is well known for its high gain^{8,10,11}, wide dynamic range^{11,12} and robust adaptation^{5,13}, mediated by receptor methylation and demethylation (by CheR and CheB).

The output of the chemotaxis pathway, the flagellar motor, is ultrasensitive to the intracellular concentration of CheY-P, with a Hill coefficient of about 10, imposing a narrow operational range for [CheY-P]⁶. Whereas precise adaptation is a robust property of the chemotaxis pathway for certain attractants, for example aspartate, the steady-state concentration of CheY-P is not⁵. Given cell-to-cell variations in the concentration of CheY-P and the fact that different cells can maintain their chemotactic sensitivity¹⁴, it has been suggested that cells might have additional molecular mechanisms to adjust the CheY-P concentration around the operational value of approximately 3 μ M⁶. One possibility is a feedback mechanism that allows a cell to adjust its kinase activity in response to motor output. This mechanism would increase the kinase activity if cells only ran, and would decrease the kinase activity if cells only tumbled. In earlier work, we looked for such a mechanism by monitoring the kinase activity with a fluorescence resonance energy transfer (FRET) technique¹⁵ while jamming flagellar bundles with an anti-filament antibody. Stopping motors had no effect on kinase activity¹⁶.

Here we report that the motor itself adapts, shifting its response function according to the steady-state concentration of CheY-P. It does this by increasing the complement of FliM when the concentration of CheY-P is low. Motor remodelling is well known for the stator elements MotA and MotB, which if defective, can be replaced by wild-type protein, as evidenced by stepwise increments in motor torque^{17,18}. Such exchange also has been visualized by total internal reflection

fluorescence (TIRF) microscopy of green fluorescent protein (GFP)-labelled protein¹⁹, and a similar technique has been used to demonstrate FliM⁹ and FliN²⁰ exchange in cells containing CheY-P. The present work addresses the functional consequences of FliM exchange. We studied *cheR cheB* cells, which are defective in methylation and demethylation, and monitored motor and kinase responses to step-addition of the non-metabolizable attractant α -methylaspartate (MeAsp), using bead²¹ and FRET¹⁵ assays. These experiments cannot be done with wild-type cells because their adaptation to aspartate is robust, so that the steady-state concentration of CheY-P does not change. Motor adaptation occurs on a minute rather than on a second timescale and does not play a direct role in sensing temporal gradients. Instead, it helps to match the operating point of the motor to the output of the chemotaxis receptor complex, obviating the requirement for fine-tuning of that output.

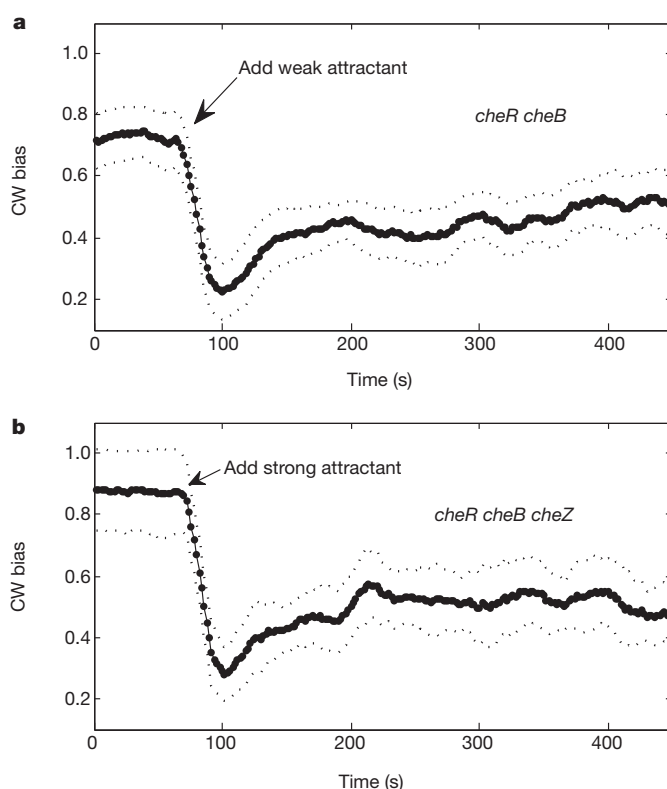


Figure 1 | Motor responses to stepwise addition of chemical attractants monitored by the bead assay. The attractants were applied at the times indicated by the arrows. Error bounds for standard errors of the mean are shown as dotted lines. **a**, Averaged responses of seven *cheR cheB* cells (JY35 carrying pKAF131) to 1 mM MeAsp (weak attractant). **b**, Averaged responses of four *cheR cheB cheZ* cells (JY32 carrying pVS7 and pKAF131) to 2 mM MeAsp + 0.5 mM L-serine (strong attractant).

¹Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Using a bead assay, we found partial adaptation in *cheR cheB* cells within 1 min following the initial response, Fig. 1a, which shows the averaged responses of seven motors on different cells to stepwise addition of 1 mM MeAsp. These results are similar to those obtained previously with tethered cells^{7,8}. A recent model suggests that partial adaptation might be due to dynamic localization of CheZ²². To test this hypothesis, we repeated the bead experiments using *cheR cheB cheZ* cells. The results were essentially the same; Fig. 1b shows the averaged responses of four motors on different cells of a *cheR cheB cheZ* strain to stepwise addition of 2 mM MeAsp + 0.5 mM L-serine (a stronger stimulus needed because of the lower sensitivity of *cheZ* strains). So CheZ is not required for this partial adaptation.

CheY-P concentrations were monitored by measuring FRET between cyan fluorescent protein (CFP)-conjugated CheZ (CheZ-CFP) and yellow fluorescent protein (YFP)-conjugated CheY (CheY-YFP). We measured responses in *cheR cheB* cells to stepwise addition of 1 mM MeAsp, Fig. 2a. The response shown in Fig. 2a is similar to that obtained previously¹¹. No adaptation is apparent. To rule out possible complications due to CheZ oligomerization²³, we also measured CFP-FliM/CheY-YFP FRET²⁴ in *cheR cheB* cells following stepwise addition of 2 mM MeAsp + 0.5 mM L-serine (a stronger stimulus needed because of the lower sensitivity of CFP-FliM/CheY-YFP FRET), as shown in Fig. 2b. No adaptation is apparent in either panel of Fig. 2, so the partial adaptation shown in Fig. 1 does not occur upstream of CheY-P. It must occur at the level of the flagellar motor.

Clockwise (CW) biases of motors were measured before addition of attractant, immediately after addition of attractant, and after time was allowed for partial adaptation. We focused on motors with pre-stimulus CW biases around 0.8 (ranging from 0.70 to 0.95). Owing to cell-to-cell variation, the lowest biases following stimulation ranged from 0 to 0.75. The concentration of CheY-P in a given cell was estimated from its CW bias at the time of the lowest bias, using the response curve measured previously⁶, shown by the red line in Fig. 3 (Hill coefficient $N_H = 10.3$, CheY-P dissociation constant $K = 3.1 \mu\text{M}$). Then the CW bias found after that cell had adapted was plotted as a

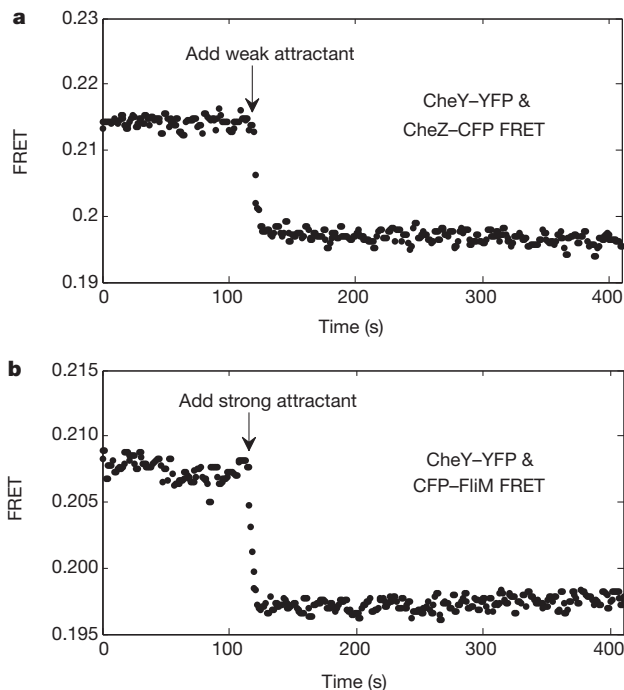


Figure 2 | FRET responses (Y/C ratio) of *cheR cheB* cells to stepwise addition of chemical attractants. The attractants were added at the times indicated by the arrows. **a**, CheY-YFP/CheZ-CFP FRET responses to 1 mM MeAsp (weak attractant). **b**, CheY-YFP/CFP-FliM FRET responses to 2 mM MeAsp + 0.5 mM L-serine (strong attractant).

function of this concentration, as shown by the blue data points in Fig. 3. The measurements were carried out for 49 motors on different *cheR cheB* cells with stepwise addition of 0.5 or 1 mM MeAsp. Following adaptation, the relationship between the CW bias and the concentration of CheY-P shifted to lower concentrations of CheY-P, increasing motor sensitivity to CheY-P.

How the motor accomplishes this shift is intriguing. We sought to explain the shift in the motor response curve by using a Monod-Wyman-Changeux (MWC) type model^{25,26}, which has been used previously to explain the motor switching kinetics²⁶. In this model, the C-ring is considered to be an allosteric switch, stochastically switching between two conformational states, counterclockwise (CCW) and CW, with N independent binding sites for CheY-P, corresponding to N units of FliM in the C-ring. The CW state has a higher affinity to CheY-P than the CCW state. The CW bias of the motor is given by $B_{CW} = (1 + [\text{CheY-P}]/K)^N / ((1 + [\text{CheY-P}]/K)^N + L(1 + [\text{CheY-P}]/(KC))^N)$, where L is the ratio of the probability that the motor is in the CCW state to the probability that it is in the CW state in the absence of CheY-P, K is the CheY-P dissociation constant for the CW state, and C is the ratio of dissociation constants for the CCW and CW states, respectively²⁶. With reasonable values for the parameters, for example, $N = 34$, $L = 10^7$ and $K = 3.1 \mu\text{M}$ ^{6,27}, the model can be fit to the ultrasensitivity data of ref. 6 with a best-fit value of C of 4.1, as shown by the red curve in Fig. 3. With these values for the parameters L , K and C , and assuming that the number of FliM units N varies with $[\text{CheY-P}]$, we can fit the data measured for the adapted motor using the MWC model with $N = N_{av} + \alpha([\text{CheY-P}] - 2.7)$, where N is written as a Taylor expansion about the average value of $[\text{CheY-P}]$, $2.7 \mu\text{M}$. We obtain a two-parameter fit with $N_{av} = 36$ and $\alpha = -1.2 \mu\text{M}^{-1}$, shown by the green curve in Fig. 3. The average value of N has increased from 34 to 36. Sensitivity of an MWC complex is known to increase with N for fixed values of L , K and C (ref. 28). Equivalently, the motor bias versus $[\text{CheY-P}]$ curve shifts to smaller $[\text{CheY-P}]$ with larger N as shown in Fig. 3. Intuitively, the fact that increasing the number of FliM units causes an increase in CW bias can be understood by considering the energetics of the switch. Each CheY-P binding decreases the energy level of CW state by a specific amount. With the values for parameters L , K and C fixed, increasing the number of FliM, that is, CheY-P binding sites, increases the number of CheY-P bound to the motor.

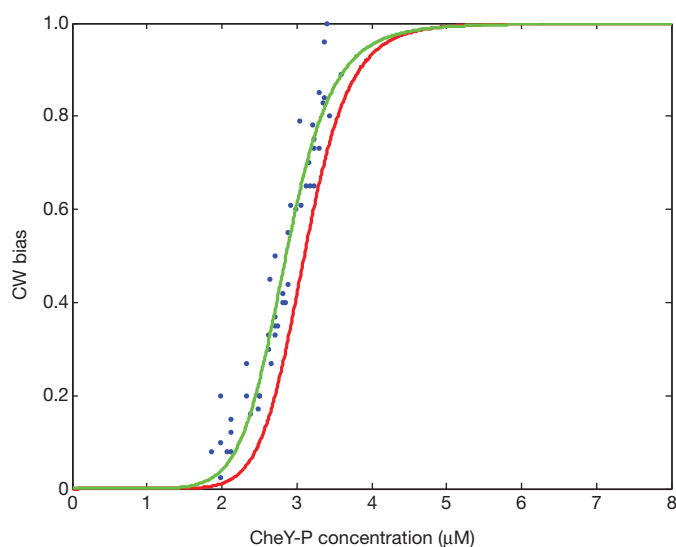


Figure 3 | CW bias as a function of CheY-P concentration. The red curve is for pre-stimulus wild-type motors, as measured in ref. 6 and fit with an MWC model with 34 FliM units²⁶. The blue dots are data for motors that have partially adapted to stepwise addition of attractant, with a two-parameter fit to the MWC model with $N_{av} = 36$ FliM units shown by the green curve; see text.

This decreases the energy level of the CW state, thereby increasing the CW bias.

To test directly for this increase of the number of FliM units, we fused YFP to the carboxy terminus of FliM and monitored the fluorescence intensity of single motors using TIRF microscopy. To minimize shifts in motor position, we tethered a *cheR cheB* strain that lacks the flagellar filament to glass via single hooks with anti-hook antibody. Changes in fluorescence were measured upon addition of 2 mM MeAsp + 0.5 mM L-serine, which should saturate the chemoreceptors and eliminate CheY-P, or of 1 mM MeAsp, which should simply reduce the concentration of CheY-P. The results are shown in Fig. 4. Figure 4a is the averaged response of 20 motors to the strong attractant, and Fig. 4b is the averaged response of 22 motors to the weak attractant. In either case, the fluorescence intensity increased following the addition of attractant on a time scale consistent with partial adaptation, by a larger amount for the stronger attractant. We compensated for fluorescence bleaching by subtracting a control curve (Fig. 4c) and fitting the results to a model in which the FliM off rate decreases when

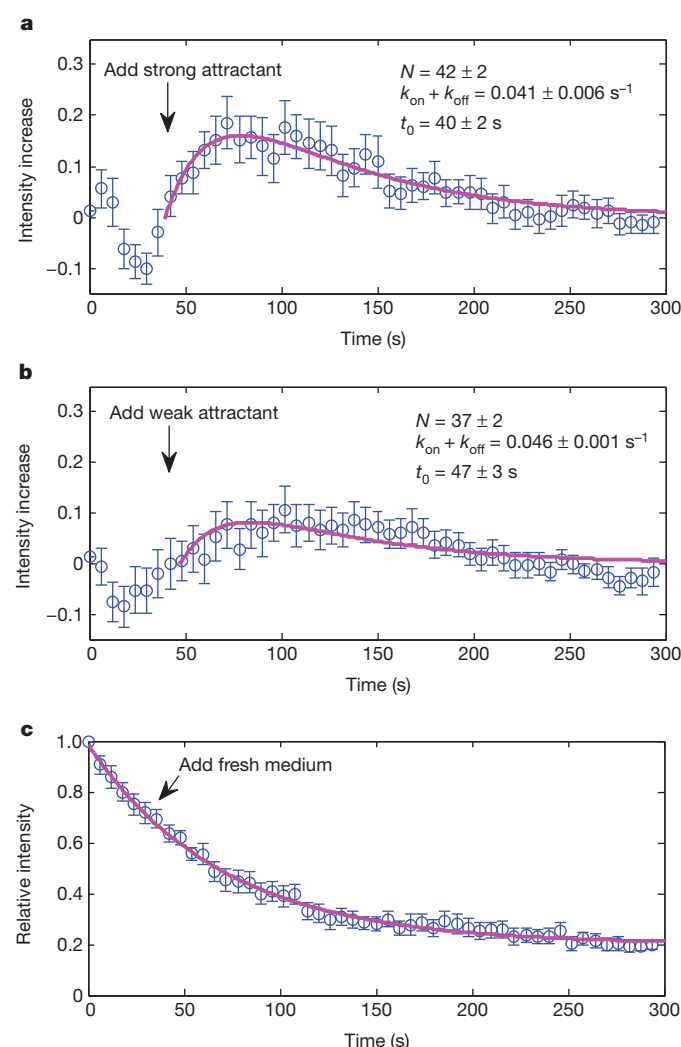


Figure 4 | Changes in single-motor FliM–YFP fluorescence intensities in *cheR cheB* cells tethered by hooks and stimulated by addition of attractant (at time t_0 , arrow). **a–c**, The intensity for a given motor was normalized by its intensity at time 0, intensities for different motors were averaged, and the fit to the control of panel c was subtracted. Fits with the model are shown in magenta and the parameters for these fits are given. Error bars are standard errors of the mean. **a**, Responses of 20 motors to addition of 2 mM MeAsp + 0.5 mM L-serine. **b**, Responses of 22 motors to addition of 1 mM MeAsp. **c**, Responses of 15 motors without addition of attractant, fit with an exponential decay function plus a constant.

the CheY-P concentration decreases; see Methods. Steady state is reached when $Nk_{\text{off}} = (M - N)k_{\text{on}}$, where N is the number of FliM molecules in the motor, and M is the maximum number of FliM binding sites in the motor. The fits are shown in magenta and the final values for N are given in each panel (assuming an initial value of 34). These values agree with those presented in Fig. 3. So the motor increases the number of FliM units as it partially adapts to a decrement in the concentration of CheY-P. By doing so, it increases its dynamic range.

We eliminated the concern that binding of CheY-P to FliM–YFP might be different than binding to wild-type FliM by using the bead assay to compare the biases, switching rates and speeds for motors of *cheR cheB* cells expressing FliM–YFP or wild-type FliM: the biases were $61 \pm 16\%$ or $58 \pm 19\%$, the switching rates were 3.8 ± 1.2 or $3.4 \pm 1.0 \text{ s}^{-1}$, and the speeds were 50.4 ± 8.4 or $51.5 \pm 8.2 \text{ Hz}$, respectively.

The motor adaptation mechanism observed here is related to the turnover of motor C-ring components discovered recently^{9,20}, where exchange of FliM was found to be signal-dependent and exhibited a similar timescale⁹. The detailed mechanism should involve changes in FliM on/off rates dependent upon either CheY-P binding or rates of motor switching. As noted earlier, the timescale for motor adaptation (1 min) is much slower than that for receptor methylation/demethylation (1 s), which enables cells to make rapid temporal comparisons; thus, motor adaptation does not play a critical role in that aspect. Instead, it helps match the operating point of the motor to the output of the chemotaxis receptor complex.

METHODS SUMMARY

All strains used in this study were derivatives of *Escherichia coli* K12 strain RP437. Cells were grown at 33 °C in 10 ml T-broth supplemented with the appropriate antibiotics and inducers to an $A_{600 \text{ nm}}$ of 0.45 to 0.50. Cells were collected by centrifugation (10 min at 1,300g), washed twice in 10 ml of motility medium (10 mM potassium phosphate/0.1 mM EDTA/1 μM methionine/10 mM lactic acid, pH 7.0), and resuspended in 10 ml of this medium. They were used immediately for experiments or stored at 4 °C for up to 2 h. All experiments were carried out with a custom-made flow chamber at room temperature.

For the bead assay, cells were sheared to truncate flagella, and 1.0- μm -diameter polystyrene latex beads were attached to the filament stubs. Rotation of the bead was monitored with a laser dark-field setup described previously²⁹. Rotational velocity as a function of time was determined for each motor and smoothed with a 25-point running average. CW bias was calculated over a 20-s interval every 2 s, leading to a plot of CW bias versus time.

FRET measurements of bacterial populations were carried out as described previously³⁰.

For TIRF measurements, cells were tethered to the bottom window of the flow chamber by single hooks using anti-FlgE antibody, following a protocol adapted from ref. 21. The fluorescence intensity of the motors was monitored with a TIRF microscope (Nikon Eclipse Ti-U), and images were recorded with a back-illuminated, cooled (–55 °C), electron-multiplying CCD camera (DV887ECS-BV, Andor Technology). Image analysis of the motor spots was carried out using a Gaussian mask method described previously¹⁹.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 November 2011; accepted 16 February 2012.

1. Sourjik, V. Receptor clustering and signal processing in *E. coli* chemotaxis. *Trends Microbiol.* **12**, 569–576 (2004).
2. Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem. Sci.* **33**, 9–19 (2008).
3. Greenfield, D. *et al.* Self-organization of the *Escherichia coli* chemotaxis network imaged with super-resolution light microscopy. *PLoS Biol.* **7**, e1000137 (2009).
4. Berg, H. C. The rotary motor of bacterial flagella. *Annu. Rev. Biochem.* **72**, 19–54 (2003).
5. Alon, U., Surette, M. G., Barkai, N. & Leibler, S. Robustness in bacterial chemotaxis. *Nature* **397**, 168–171 (1999).
6. Cluzel, P., Surette, M. & Leibler, S. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science* **287**, 1652–1655 (2000).
7. Stock, J., Kersulis, G. & Koshland, D. E. Jr. Neither methylating nor demethylating enzymes are required for bacterial chemotaxis. *Cell* **42**, 683–690 (1985).

8. Segall, J. E., Block, S. M. & Berg, H. C. Temporal comparisons in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA* **83**, 8987–8991 (1986).
9. Delalez, N. J. *et al.* Signal-dependent turnover of the bacterial flagellar switch protein FliM. *Proc. Natl Acad. Sci. USA* **107**, 11347–11351 (2010).
10. Duke, T. A. J. & Bray, D. Heightened sensitivity of a lattice of membrane of receptors. *Proc. Natl Acad. Sci. USA* **96**, 10104–10108 (1999).
11. Sourjik, V. & Berg, H. C. Receptor sensitivity in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA* **99**, 123–127 (2002).
12. Mello, B. A. & Tu, Y. Effects of adaptation in maintaining high sensitivity over a wide dynamic range of backgrounds for *Escherichia coli* chemotaxis. *Biophys. J.* **92**, 2329–2337 (2007).
13. Hansen, C. H., Endres, R. G. & Wingreen, N. S. Chemotaxis in *Escherichia coli*: a molecular model for robust precise adaptation. *PLoS Comput. Biol.* **4**, e1 (2008).
14. Levin, M. D., Morton-Firth, C. J., Abouhamad, W. N., Bourret, R. B. & Bray, D. Origins of individual swimming behavior in bacteria. *Biophys. J.* **74**, 175–181 (1998).
15. Sourjik, V., Vaknin, A., Shimizu, T. S. & Berg, H. C. *In vivo* measurement by FRET of pathway activity in bacterial chemotaxis. *Methods Enzymol.* **423**, 365–391 (2007).
16. Shimizu, T. S., Delalez, N., Pichler, K. & Berg, H. C. Monitoring bacterial chemotaxis by using bioluminescence resonance energy transfer: absence of feedback from the flagellar motors. *Proc. Natl Acad. Sci. USA* **103**, 2093–2097 (2006).
17. Blair, D. F. & Berg, H. C. Restoration of torque in defective flagellar motors. *Science* **242**, 1678–1681 (1988).
18. Reid, S. W. *et al.* The maximum number of torque-generating units in the flagellar motor of *Escherichia coli* is at least 11. *Proc. Natl Acad. Sci. USA* **103**, 8066–8071 (2006).
19. Leake, M. C. *et al.* Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature* **443**, 355–358 (2006).
20. Fukuoka, H., Inoue, Y., Terasawa, S., Takahashi, H. & Ishijima, A. Exchange of rotor components in functioning bacterial flagellar motor. *Biochem. Biophys. Res. Commun.* **394**, 130–135 (2010).
21. Yuan, J. & Berg, H. C. Resurrection of the flagellar motor near zero load. *Proc. Natl Acad. Sci. USA* **105**, 1182–1185 (2008).
22. Lipkow, K. Changing cellular localization of CheZ predicted by molecular simulations. *PLoS Comput. Biol.* **2**, e39 (2006).
23. Blat, Y. & Eisenbach, M. Oligomerization of the phosphatase CheZ upon interaction with the phosphorylated form of CheY, the signal protein of bacterial chemotaxis. *J. Biol. Chem.* **271**, 1226–1231 (1996).
24. Sourjik, V. & Berg, H. C. Binding of the *Escherichia coli* response regulator CheY to its target measured *in vivo* by fluorescence resonance energy transfer. *Proc. Natl Acad. Sci. USA* **99**, 12669–12674 (2002).
25. Monod, J., Wyman, J. & Changeux, J. P. On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118 (1965).
26. Alon, U. *et al.* Response regulator output in bacterial chemotaxis. *EMBO J.* **17**, 4238–4248 (1998).
27. Turner, L., Caplan, S. R. & Berg, H. C. Temperature induced switching of the bacterial flagellar motor. *Biophys. J.* **71**, 2227–2233 (1996).
28. Sourjik, V. & Berg, H. C. Functional interactions between receptors in bacterial chemotaxis. *Nature* **428**, 437–441 (2004).
29. Yuan, J., Fahrner, K. A. & Berg, H. C. Switching of the bacterial flagellar motor near zero load. *J. Mol. Biol.* **390**, 394–400 (2009).
30. Shimizu, T. S., Tu, Y. & Berg, H. C. A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli. *Mol. Syst. Biol.* **6**, 382 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by National Institutes of Health Grant AI016478. R.W.B. is a recipient of an EMBO Long-Term Fellowship.

Author Contributions J.Y. and H.C.B. planned the work and wrote the first draft of the paper. J.Y. performed the research with help on the MWC model and TIRF experiment and analysis from R.W.B. and on TIRF analysis from B.G.H.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.C.B. (hberg@mcb.harvard.edu).

METHODS

Strains and plasmids. All strains used in this study are derivatives of *E. coli* K12 strain RP437 (ref. 31): JY32 (*cheR cheB cheY cheZ fliC*), JY35 (*cheR cheB fliC*), RP2893 ($\Delta 2206(\text{tap-cheZ})$)¹¹, JY37 (*cheR cheB cheY fliM*), and JY40 (*cheR cheB fliM fliC*). The *fliM-eyfp*^{A206K} fusion with a 3 × glycine linker was cloned into pTrc99A³² under an isopropyl-β-D-thiogalactoside (IPTG)-inducible promoter, yielding pRWB7. pDFB72 carrying wild-type *fliM* on pTrc99A was a gift from D. Blair. pVS7 carrying wild-type *cheY* on a pBAD18-Kan³³ vector was a gift from V. Sourjik. pVS18 carrying *cheY-eyfp* on pTrc99A, pVS31 carrying *ecfp-fliM* on pBAD33 (ref. 33), and pVS54 carrying *cheZ-ecfp* on pBAD33, were described previously^{11,24}. pKAF131 carrying the sticky *fliC* allele under control of the native *fliC* promoter, was described previously³⁴. For studies of *cheR cheB* cells with the bead assay, JY35 carrying pKAF131 was used. For studies of *cheR cheB cheZ* cells with the bead assay, JY32 carrying pVS7 and pKAF131 was used. For CheZ–CFP/CheY–YFP FRET studies of *cheR cheB* cells, RP2893 carrying pVS18 and pVS54 was used. For CFP–FliM/CheY–YFP FRET studies of *cheR cheB* cells, JY37 carrying pVS18 and pVS31 was used. For the TIRF studies of single motors, JY40 carrying pRWB7 was used. For comparison of motors with wild-type FliM and FliM–YFP, JY40 carrying pDFB72 and pKAF131, and JY40 carrying pRWB7 and pKAF131 were used. Cells were grown at 33 °C in 10 ml T-broth (1% tryptone and 0.5% NaCl) supplemented with the appropriate antibiotics (ampicillin: 100 µg ml^{−1}, kanamycin: 50 µg ml^{−1}, chloramphenicol: 34 µg ml^{−1}) and inducers (0.005% arabinose for the bead assay, 0.01% arabinose and 50 µM IPTG for the FRET studies, 100 µM IPTG for the TIRF studies) to an $A_{600\text{ nm}}$ of 0.45 to 0.50. Cells were collected by centrifugation (10 min at 1,300g), washed twice in 10 ml of motility medium (10 mM potassium phosphate/0.1 mM EDTA/1 µM methionine/10 mM lactic acid, pH 7.0), and resuspended in 10 ml of this medium. They were used immediately for experiments or stored at 4 °C for up to 2 h.

Bead assay and data analysis. Cells were sheared to truncate flagella by passing 1 ml of the washed-cell suspension 50 times between two syringes equipped with 23-gauge needles and connected by a 7-cm length of polyethylene tubing (0.58 mm internal diameter, catalogue no. 427411; Becton Dickinson). The sheared cell suspension was centrifuged and resuspended in 0.5 ml of motility medium. 50 µl of this suspension was placed on a glass coverslip coated with poly-L-lysine (0.01%, catalogue no. P4707; Sigma) and allowed to stand for 2 min, then 5 µl of 1.0-µm-diameter polystyrene latex beads (2.69%, catalogue no. 07310; Polysciences) was added, mixed by gentle pipetting, and allowed to stand for 2 min. The coverslip was installed as the top window of a flow chamber³⁵ and rinsed with motility medium. The chamber was kept under a constant flow of buffer (400 µl min^{−1}) by a syringe pump (Harvard Apparatus). Rotation of the bead was monitored with a laser dark-field setup described previously²⁹. Outputs from the photomultiplier tubes were directly coupled to an eight-pole low-pass Bessel filter (3384, Krohn-Hite) with a cutoff frequency of 200 Hz and sampled at 500 Hz using LabView. For each experiment, the rotation of the bead was monitored for 70 s, then the medium was switched to attractants, and the rotation was monitored further for about 400 s. Rotational velocity as a function of time was determined for each motor as described previously²⁹ and smoothed with a 25-point running average. CW bias was calculated over a 20-s interval every 2 s, leading to a plot of CW bias versus time.

FRET measurements. FRET measurements of bacterial populations were carried out as described previously^{15,30}, except that the epifluorescent illumination was provided by a LED white light source (MCWHL2-C3, Thorlabs) through an excitation bandpass filter (FF01-438/24-25, Semrock). For each experiment, 4 ml of the washed-cell suspension was centrifuged and resuspended in 55 µl of motility medium, which was placed on a glass coverslip coated with poly-L-lysine (0.1%, catalogue no. P8920, Sigma) and allowed to stand for 5 min. The coverslip was installed as the top window of a flow chamber³⁵ and rinsed with motility medium. The chamber was kept under a constant flow of buffer (500 µl min^{−1}). Epifluorescent emission was split into donor (cyan, C) and acceptor (yellow, Y) channels and collected by photon-counting photomultipliers (H7421-40, Hamamatsu). Signal intensities of these two channels were recorded by a computer running LabView, and the ratio between them ($R = Y/C$) provided an indicator of FRET activity. The FRET traces were smoothed with a median filter of rank 3.

TIRF measurements, data analysis and fits to the model. Cells were tethered to glass by hooks using anti-FlgE antibody, following a protocol adapted from ref. 21: 350 µl of the washed-cell suspension was centrifuged and resuspended in 100 µl of motility medium; 10 µl of anti-FlgE antibody (0.1 mg ml^{−1}) was added, and the mixture was incubated at 23 °C for 25 min. The antibody-treated cells were washed twice with 350 µl of motility medium and gently resuspended in 55 µl of motility medium. This cell suspension was placed on the bottom coverslip of a flow chamber

(which was washed earlier with ethanol and distilled H₂O and air-dried for 2 h) and allowed to stand for 15 min. The top coverslip of the flow chamber was then installed and rinsed with motility medium. The chamber was kept under a constant flow of buffer (100 µl min^{−1}). For each experiment, the flow was switched to attractants at 1 min before the start. It took about 1 min and 30 s for the attractant to reach the flow chamber and about 30 s to replace the medium, so effectively the medium reached the cells at between 30 s to 60 s after the start of each experiment. Only stably rotating and switching motors were monitored. The motors observed usually started with high CW bias; upon addition of strong attractant, they changed to exclusively CCW and remained 100% CCW throughout the observation time; upon addition of weak attractant, their CW bias reduced and later partially recovered; upon addition of motility medium, their CW bias did not change. The fluorescent intensity of the motors was monitored with a TIRF microscope (Nikon Eclipse Ti-U), and images were recorded at 65 nm per pixel with a back-illuminated, cooled (−55 °C), electron-multiplying CCD camera (DV887ECS-BV, Andor Technology). The camera was controlled by Andor Solis software running on a desktop computer. Image acquisition was under Andor Solis 'kinetic' mode, with 200 ms exposures every 6 s for 50 exposures for each motor. The laser illumination was blocked between exposures.

Images of the motor spots showed radially symmetric and approximately Gaussian intensity profiles. The width of these spots was about 5 pixels (325 nm). The fluorescent intensity centroid for each motor was calculated using a Gaussian mask method described previously^{19,36}. Specifically, an initial estimate was made based on the peak pixel intensity, a 9 × 9 pixel region of interest (ROI) was defined centring on the initial motor centroid, and the motor centroid was calculated as follows: first, a circular motor mask of diameter 300 nm was applied to the ROI centring on the current motor centroid. Second, pixel intensities within the motor mask were multiplied by a radially symmetric two-dimensional Gaussian mask of fixed half-width 170 nm, and a revised estimate for the motor centroid was calculated using a weighted average. Lastly, the previous two steps were iterated either 150 times or until the motor mask began clipping the side of the ROI. We also calculated the motor centroid with a two-dimensional Gaussian fitting, and both methods yielded comparable results. After the centroid was calculated, the background intensity was defined as the mean pixel intensity within the ROI but external to the final motor mask, and the motor intensity was calculated as the sum of all pixel intensities within the motor mask after subtraction of the background intensity from each pixel value.

The model assumes that CheY-P binding destabilizes FliM, so that when [CheY-P] suddenly decreases due to addition of attractant, k_{off} (the off rate of each FliM unit) decreases, while k_{on} remains the same. When the number of FliM units (N) in the C-ring reaches a new steady state, $Nk_{\text{off}} = (M - N)k_{\text{on}}$, where M is the maximum number of FliM binding sites in a motor. The pre-stimulus N is assumed to be 34. During the response to the attractant step, the increment of N satisfies $dn = ((M - (n + 34))k_{\text{on}} - (n + 34)k_{\text{off}})dt$, while the increment of the normalized motor intensity satisfies $df = dn/a - \lambda df dt$, where a is the normalization factor that converts the number of FliM units to fluorescence intensity, and λ is the fluorescence bleaching rate obtained by fitting the control curve (Fig. 4c). Solving these two differential equations with the initial conditions: $n(0) = 0$, $f(0) = 0$ leads to:

$$f = \frac{(k_{\text{on}}M - 34(k_{\text{on}} + k_{\text{off}}))}{a(k_{\text{on}} + k_{\text{off}} - \lambda)}(e^{-\lambda t} - e^{-(k_{\text{on}} + k_{\text{off}})t}) \\ = \frac{(N - 34)(k_{\text{on}} + k_{\text{off}})}{a(k_{\text{on}} + k_{\text{off}} - \lambda)}(e^{-\lambda t} - e^{-(k_{\text{on}} + k_{\text{off}})t})$$

If the time of arrival of the attractant at the cell is t_0 instead of 0, change t in the above equations to $t - t_0$.

- Parkinson, J. S. Complementation analysis and deletion mapping of *Escherichia coli* mutants defective in chemotaxis. *J. Bacteriol.* **135**, 45–53 (1978).
- Amann, E., Ochs, B. & Abel, K.-J. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**, 301–315 (1988).
- Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
- Yuan, J., Fahrner, K. A., Turner, L. & Berg, H. C. Asymmetry in the clockwise and counter-clockwise rotation of the bacterial flagellar motor. *Proc. Natl Acad. Sci. USA* **107**, 12846–12849 (2010).
- Berg, H. C. & Block, S. M. A miniature flow cell designed for rapid exchange of media under high-power microscope objectives. *J. Gen. Microbiol.* **130**, 2915–2920 (1984).
- Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775–2783 (2002).

Trans-synaptic Teneurin signalling in neuromuscular synapse organization and target choice

Timothy J. Mosca^{1*}, Weizhe Hong^{1*}, Vardhan S. Dani¹, Vincenzo Favaloro¹ & Liqun Luo¹

Synapse assembly requires trans-synaptic signals between the pre- and postsynapse¹, but our understanding of the essential organizational molecules involved in this process remains incomplete². Teneurin proteins are conserved, epidermal growth factor (EGF)-repeat-containing transmembrane proteins with large extracellular domains³. Here we show that two *Drosophila* Teneurins, Ten-m and Ten-a, are required for neuromuscular synapse organization and target selection. Ten-a is presynaptic whereas Ten-m is mostly postsynaptic; neuronal Ten-a and muscle Ten-m form a complex *in vivo*. Pre- or postsynaptic Teneurin perturbations cause severe synapse loss and impair many facets of organization trans-synaptically and cell autonomously. These include defects in active zone apposition, release sites, membrane and vesicle organization, and synaptic transmission. Moreover, the presynaptic microtubule and postsynaptic spectrin cytoskeletons are severely disrupted, suggesting a mechanism whereby Teneurins organize the cytoskeleton, which in turn affects other aspects of synapse development. Supporting this, Ten-m physically interacts with α -Spectrin. Genetic analyses of *teneurin* and *neuroligin* reveal that they have differential roles that synergize to promote synapse assembly. Finally, at elevated endogenous levels, Ten-m regulates target selection between specific motor neurons and muscles. Our study identifies the Teneurins as a key bi-directional trans-synaptic signal involved in general synapse organization, and demonstrates that proteins such as these can also regulate target selection.

Vertebrate teneurins are enriched in the developing brain^{4,5}, localize to synapses in culture⁶, and pattern visual connections⁷. Both *Drosophila* Teneurins, Ten-m and Ten-a, function in olfactory synaptic partner matching⁸ and were further identified in neuromuscular junction (NMJ) defect screens^{9,10}, with Ten-m also affecting motor axon guidance¹¹. We examine their roles and the underlying mechanisms involved in synapse development.

Both Ten-m and Ten-a were enriched at the larval NMJ (Fig. 1a and Supplementary Fig. 1a). Ten-a was detected at neuronal membranes; this staining was undetectable beyond background in *ten-a* null mutants (Supplementary Fig. 1b) and barely detectable after neuronal *ten-a* RNA interference (RNAi; Supplementary Fig. 1c), indicating that Ten-a is predominantly presynaptic. Partial co-localization was observed between Ten-a and the periaxial zone marker Fasciclin 2 (ref. 12) as well as the active zone marker Bruchpilot¹³ (Fig. 1b, c), suggesting that Ten-a is localized to the junction between the periaxial zone and the active zone. Ten-m appeared strongly postsynaptic and surrounded each bouton (Fig. 1a and Supplementary Fig. 1a, d). Muscle-specific *ten-m* RNAi eliminated the postsynaptic staining, but uncovered weak presynaptic staining (Supplementary Fig. 1e) that ubiquitous *ten-m* RNAi eliminated (Supplementary Fig. 1f). Thus, the Ten-m signal was specific and, although partly presynaptic, enriched postsynaptically. Consistently, muscle Ten-m colocalized extensively with Dlg (Fig. 1d) and completely with α -Spectrin (Fig. 1e) and is thus probably coincident with all postsynaptic membranes.

The localization of Ten-a and Ten-m suggested their trans-synaptic interaction. To examine this, we co-expressed Myc-tagged Ten-a in nerves using the Q system¹⁴ and haemagglutinin (HA)-tagged Ten-m in muscles using GAL4. Muscle Ten-m was able to co-immunoprecipitate nerve Ten-a from larval synaptosomes (Fig. 1f), suggesting that the Teneurins form a heterophilic trans-synaptic receptor pair at the NMJ.

To determine Teneurin function at the NMJ, we examined the *ten-a* null allele and larvae with neuron or muscle RNAi of *ten-a* and/or *ten-m*. Following such perturbations, bouton number and size were altered: the quantity was reduced by 55% (Fig. 2a–c, g and Supplementary Fig. 2)

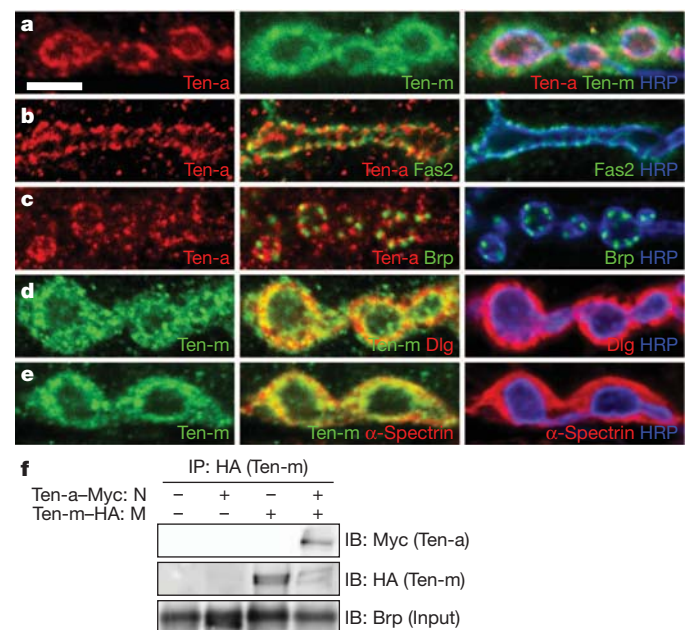


Figure 1 | Teneurins are enriched at and interact across *Drosophila* neuromuscular synapses. a–e, Representative single confocal sections of synaptic boutons stained with antibodies against Ten-a (red) or Ten-m (green), horseradish peroxidase (HRP) to mark the neuronal membrane (blue), and a synaptic marker as indicated. a, Ten-a is associated with presynaptic membranes and Ten-m largely with the surrounding postsynapse (a). b, c, Ten-a shows limited co-localization with the periaxial zone marker Fasciclin 2 (b), and Bruchpilot (Brp), an active zone marker (c). d, e, Ten-m co-localizes with, and extends beyond, Dlg (d) and completely co-localizes with muscle α -Spectrin (e). f, Immunoblots (IB) of larval synaptosomes expressing neuronal Flag-Myc-tagged Ten-a (N) and muscle Flag-HA-tagged Ten-m (M) and immunoprecipitated (IP) using antibodies to HA. Ten-a is detected in the pull-down, indicating that nerve Ten-a and muscle Ten-m interact across the NMJ. This is not seen in control lanes. Owing to low expression, neither transgene product is detectable in input lysates, which are enriched in Brp. Scale bar, 5 μ m.

¹Department of Biology, Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA.

*These authors contributed equally to this work.

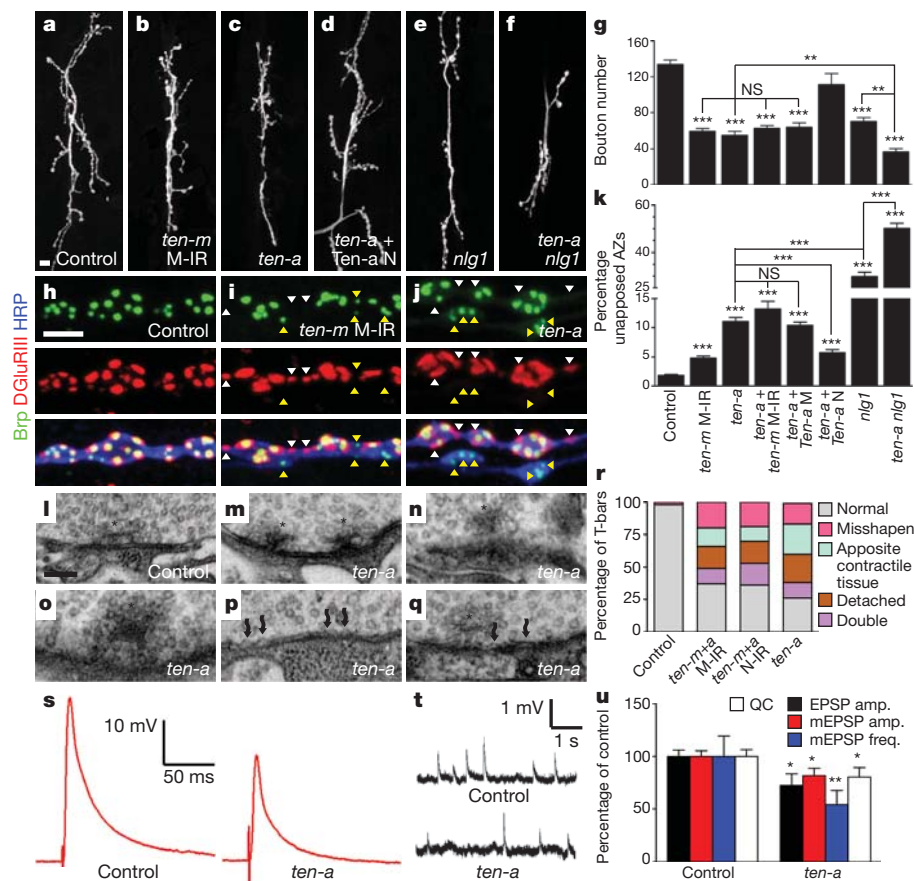


Figure 2 | Teneurins affect the structure and function of the neuromuscular synapse. **a–f**, Representative NMJs stained with antibodies to HRP. Muscle-specific *ten-m* RNAi (M-IR) and lots of *ten-a* decrease bouton number. Neuronal (Ten-a N) but not muscle (Ten-a M) restoration of *ten-a* expression rescues this phenotype. These defects resemble *nlg1* mutants and are enhanced in *ten-a nlg1* double mutants. **g**, Quantification of bouton number. **h–j**, Representative NMJs stained with antibodies to Brp (green), the glutamate receptor subunit DGluRIII (red) and HRP (blue). In control larvae (**h**), Brp and DGluRIII puncta properly appose. *teneurin* perturbations (**i**, **j**) disrupt this active zone (yellow arrowhead) and glutamate receptor apposition (white arrowhead). **k**, Quantification of unapposed active zone/glutamate receptor pairs. For each quantification, $n \geq 8$ larvae, 16 NMJs. **l–q**, Transmission electron microscopy of active zone T-bars (asterisks) in control larvae (**l**) and

ten-a mutants (**m–q**) showing double T-bars (**m**), detached T-bars (**n**), misshapen T-bars (**o**), membrane ruffling (**p**, waved arrows) and T-bars facing contractile tissue (**q**). Some images show multiple defects. **r**, Distribution of T-bar defects as a percentage of the total T-bars. N-IR, neuronal RNAi. For each genotype, $n \geq 3$ larvae, 40 boutons. **s**, **t**, Representative evoked EPSP (**s**) and mEPSP (**t**) traces from control and *ten-a* mutant genotypes. **u**, Quantification of mean EPSP amplitude (black), mEPSP amplitude (red), mEPSP frequency (blue) and quantal content (QC, white) expressed as a percentage of the control average. For all genotypes, $n \geq 7$ larvae, 8 muscles. Error bars represent s.e.m. Scale bars, 10 μ m (**a–f**), 5 μ m (**h–j**), 100 nm (**l–q**). *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, NS, not significant. Statistical comparisons are with control unless noted.

and the incidence of large boutons markedly increased (Supplementary Fig. 2k). Both changes indicate impaired synaptic morphogenesis. The reduction in bouton number was probably cumulative through development, as it was visible in first instar *ten-a* mutants and persisted (Supplementary Fig. 2k). In the *ten-a* mutant, bouton morphogenesis was rescued by restoring Ten-a expression in neurons, but not muscles (Fig. 2d, g and Supplementary Fig. 2). Neuronal Ten-m overexpression could not substitute for the lack of Ten-a, revealing their non-equivalence (Supplementary Fig. 2e, l). Neuronal knockdown of Ten-a or Ten-m resulted in fewer synaptic boutons (Supplementary Fig. 2f–h, l), indicating that both have a presynaptic function, although presynaptic Ten-a has a more predominant role (Supplementary Fig. 2l). Moreover, knocking down postsynaptic Ten-m in the *ten-a* mutant did not enhance the phenotype (Fig. 2g). Thus, presynaptic Ten-a (and, to a lesser extent, Ten-m) and postsynaptic Ten-m are required for synapse development.

teneurin perturbation also caused defects in the apposition between presynaptic active zones (release sites) and postsynaptic glutamate receptor clusters¹⁵ (Fig. 2h and Supplementary Fig. 3): up to 15% of the active zones/receptor clusters lacked their partner compared to 1.8% in controls (Fig. 2h–k). Under electron microscopy, active zones

are marked by electron-dense membranes and single presynaptic specializations called T-bars (Fig. 2l), which enable synapse assembly, vesicle release and Ca^{2+} -channel clustering¹⁶. Teneurin disruption caused defects (Fig. 2m–r and Supplementary Fig. 3) in T-bar ultrastructure (Fig. 2m–o), membrane organization, and apposition to contractile tissue (Fig. 2p, q). Teneurin perturbation also impaired postsynaptic densities while increasing membrane ruffling (Supplementary Table 1), further indicating organizational deficiency. These phenotypes resemble mutants with adhesion and T-bar biogenesis defects^{17,18}, suggesting a role for Teneurins in synaptic adhesion and stability. Synaptic vesicle populations similarly required Teneurins for clustering at the bouton perimeter and proper density (Supplementary Fig. 4). As these effects are not synonymous with active zone disruption¹⁹, Teneurins are also required for synaptic vesicle organization.

Synapses lacking *teneurin* were also functionally impaired. The mean amplitude of evoked excitatory postsynaptic potentials (EPSPs) in larvae was decreased by 28% in the *ten-a* mutant (Fig. 2s, u). Spontaneous miniature EPSPs showed a 20% decrease in amplitude, a 46% decrease in frequency (Fig. 2t, u), and an altered amplitude distribution compared with control (Supplementary Fig. 5a). These defects resulted in a 20% reduction in quantal content (Fig. 2u), which could be partly due to

fewer boutons and release sites. However, release probability may also be reduced, as suggested by an increased paired pulse ratio in *ten-a* mutants (Supplementary Fig. 5d, e). The decay kinetics of responses were faster in *ten-a* mutants, suggesting additional postsynaptic effects on glutamate receptors and/or intrinsic membrane properties (Supplementary Fig. 5b, c). Further, FM1-43 dye loading revealed markedly defective vesicle cycling in *ten-a* mutants (Supplementary Fig. 5f, h). Consistent with physiological impairment, *teneurin*-perturbed larvae exhibited profound locomotor defects (Supplementary Fig. 5i). In summary, Teneurins are required for multiple aspects of NMJ organization and function.

As a potential mechanism for synaptic disorganization following *teneurin* perturbation, we examined the pre- and postsynaptic cytoskeletons. In the presynaptic terminal, organized microtubules contain Futsch (a microtubule-binding protein)-positive 'loops', whereas disorganized microtubules possess punctate, 'unbundled' Futsch²⁰. Each classification normally represented ~10% (often distal) of boutons (Fig. 3a, d and Supplementary Fig. 6). Upon *teneurin* perturbation, many more boutons had unbundled Futsch (Fig. 3b, c and Supplementary Fig. 6) whereas those with looped microtubules were decreased by 62–95% (Fig. 3d). Therefore, proper microtubule organization requires pre- and postsynaptic Teneurins. In contrast to mild active zone/glutamate receptor apposition defects, most boutons displayed microtubule organizational defects.

teneurin perturbation also severely disrupted the postsynaptic spectrin cytoskeleton, with which Ten-m colocalized (Fig. 1e). Postsynaptic α -Spectrin normally surrounds the bouton (Fig. 3e). Perturbing neuronal or muscle Teneurins markedly reduced postsynaptic α -Spectrin without affecting Dlg (Fig. 3f–h and Supplementary Fig. 7). Postsynaptic β -Spectrin²¹, Adducin²² and Wsp were similarly affected (Supplementary Fig. 8). In muscle, α -Spectrin is coincident with and essential for the integrity of the membranous subsynaptic reticulum (SSR)^{21,23}. Consistent with this, *teneurin* disruption reduced SSR width up to 70% (Supplementary Fig. 9d–g) and increased the frequency of 'ghost' boutons, which are failures of postsynaptic membrane organization²³

(Supplementary Fig. 9a–d). Thus, Teneurins are involved in the organization of the pre- and postsynaptic cytoskeletons and postsynaptic membranes. Further, endogenous α -Spectrin co-immunoprecipitated with muscle-expressed, Flag-tagged Ten-m (Fig. 3i), suggesting that Ten-m physically links the synaptic membrane to the cytoskeleton.

Because the most severe defects following *teneurin* perturbation were cytoskeletal, we propose that Teneurins primarily organize the presynaptic microtubule and postsynaptic spectrin-based cytoskeletons (Fig. 3j), which then organize additional synaptic aspects^{20,21}. However, such a solitary role cannot fully explain the observed phenotypes. The reduction in bouton number associated with cytoskeletal disruption is milder than that following *teneurin* disruption^{20,21,24}. Also, although active zone dynamics are affected by cytoskeletal perturbation²¹, defects in apposition are not^{21,25}. Moreover, the T-bar structural defects more closely resemble synapse adhesion and active zone formation defects^{17,18}. Thus, Teneurins may regulate release site organization and synaptic adhesion independent of the cytoskeleton (Fig. 3j).

Our data also indicate that Teneurins act bi-directionally across the synaptic cleft. Ten-a acts predominantly in neurons, as evidenced by localization, phenotypes caused by neuronal (but not muscle) knock-down, and mutant rescue by neuronal (but not muscle) expression (Figs 2 and 3 and Supplementary Figs 2–4, 6, 7 and 9). Yet, in addition to the presynaptic phenotypes, many others were postsynaptic, including reduced muscle spectrin, SSR, and membrane apposition (Fig. 3 and Supplementary Figs 7–9). Similarly, although Ten-m is present both pre- and postsynaptically, muscle knockdown resulted in pre-synaptic defects, including microtubule and vesicle disorganization, reduced active zone apposition, and T-bar defects (Figs 2 and 3 and Supplementary Figs 3, 4, 6 and 7). Thus, Teneurins function in bi-directional trans-synaptic signalling to organize neuromuscular synapses. This may involve downstream pathways or simply establish an organizational framework by the receptors themselves. Moreover, as the results of single disruptions of neuronal *ten-a* or muscle *ten-m* were similarly severe and not enhanced by combination (Figs 2g and 3d, h and Supplementary Fig. 2k), both Ten-a and Ten-m probably

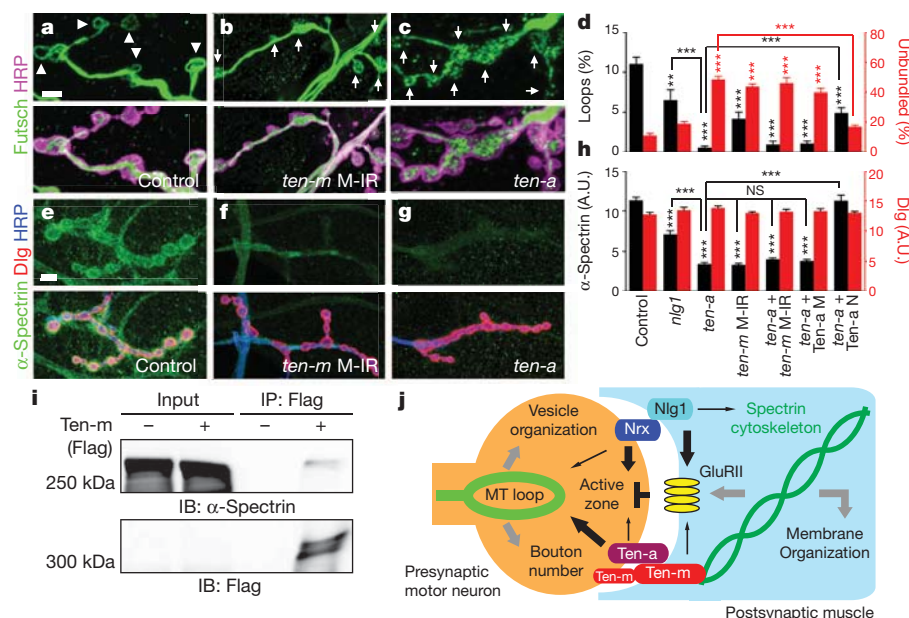


Figure 3 | Teneurin perturbation results in marked cytoskeletal disorganization. a–c, Representative NMJs stained with antibodies to Futsch (green) and HRP (magenta). Arrowheads indicate looped organization. Arrows indicate unbundled Futsch. d, Quantification of the percentage of total boutons with looped or unbundled microtubules. e–g, Representative NMJs stained with antibodies to α -Spectrin (green), Dlg (red) and HRP (blue). Following *teneurin* perturbation, α -Spectrin staining is largely lost. Axonal α -Spectrin is unaffected by muscle *teneurin* RNAi (f). h, Quantification of α -Spectrin (green)

and Dlg (red) fluorescence. A.U., arbitrary units. For all genotypes, $n \geq 6$ larvae, 12 NMJs. i, Immunoblots (IB) showing that α -Spectrin is detected in the Flag immunoprecipitates (IP) of larvae expressing muscle Flag-HA-tagged Ten-m but not in control larvae. Owing to low expression, Flag-HA-Ten-m is only detectable after enrichment by immunoprecipitation. j, Model showing the roles of Teneurins, Neurexin and Neurotrophin at the NMJ. Arrow size represents the relative contribution of each pathway to the cellular process as inferred from mutant phenotypic severity. Scale bars, 5 μ m. *** $P < 0.001$, NS, not significant.

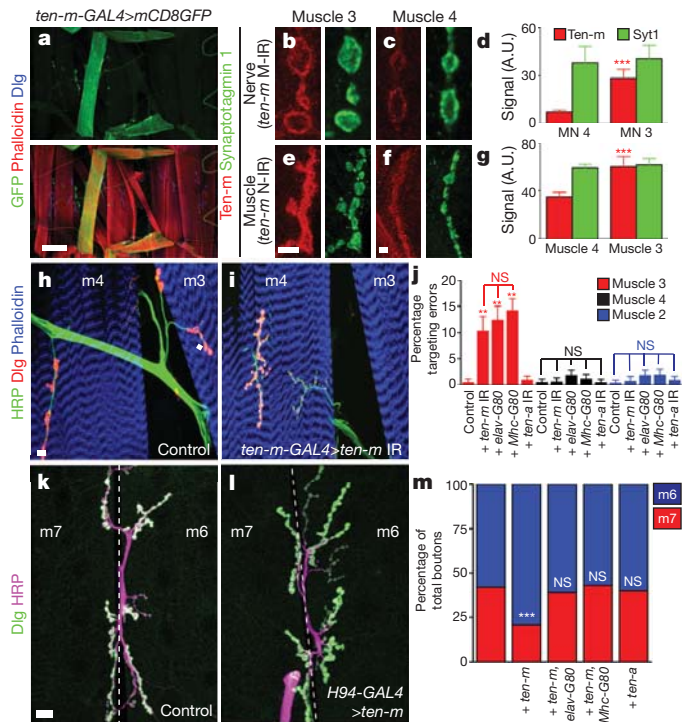
function in the same pathway. Our finding that Ten-a and Ten-m co-immunoprecipitate from different cells *in vitro*⁸ and across the NMJ *in vivo* (Fig. 1f) further suggests a signal via a trans-synaptic complex. Teneurin function, however, may not be solely trans-synaptic. In some cases (vesicle density, SSR width), cell-autonomous knockdown resulted in stronger phenotypes than knocking down in synaptic partners (Supplementary Figs 3, 4, 9 and Supplementary Table 1). This suggests additional cell-autonomous roles in addition to trans-synaptic Teneurin signalling.

Signalling involving the transmembrane proteins Neurexin and Neuroligin also mediates synapse development²⁶. In *Drosophila*, Neurexin (*nrx*) and Neuroligin 1 (*nlg1*) mutations cause phenotypes similar to *teneurin* perturbation: reductions in bouton number, active zone organization, transmission, and SSR width^{27,28}. *nlg1* and *nrx* mutations do not enhance each other, suggesting that they function in the same pathway²⁸. Consistently^{27,28}, we found that *nrx* and *nlg1* mutants exhibited largely similar phenotypes (data not shown). To investigate the relationship between the *teneurin* genes and *nrx* and *nlg1*, we focused on the *nlg1* null mutant. Both Nlg1 tagged with enhanced green fluorescent protein (Nlg1-eGFP) and endogenous Ten-m occupied a similar postsynaptic space (Supplementary Fig. 10a). *teneurin* and *nlg1* loss-of-function mutations also displayed similar bouton number reductions (Fig. 2e, g), vesicle disorganization (Supplementary Fig. 4), and ghost bouton frequencies (Supplementary Fig. 9). Other phenotypes showed notable differences in severity. In *nlg1* mutants, there was a 29% failure of active zone/glutamate receptor apposition (Fig. 2k and Supplementary Fig. 10d), compared to 15% for the strongest *teneurin* perturbation. The cytoskeleton of *nlg1* mutants, however, was only mildly impaired compared to that seen with *teneurin* perturbations (Fig. 3d, h and Supplementary Figs 6 and 7).

To examine further the interplay of *teneurin* and *nlg1*, we analysed *ten-a nlg1* double mutants. Both single mutants were viable, despite their synaptic defects. Double mutants, however, were larval lethal. We obtained rare escapers, which showed a 72% reduction in boutons, compared to a 50–55% decrease in single mutants (Fig. 2e). Active zone apposition in double mutants was enhanced synergistically over either single mutant (Fig. 2k and Supplementary Fig. 10e). Cytoskeletal defects in the double mutant resembled the *ten-a* mutant (Fig. 3 and Supplementary Figs 6 and 7). These data suggest that *teneurin* genes and *nrx* and *nlg1* act in partially overlapping pathways, cooperating to organize synapses properly, with Teneurins contributing more to cytoskeletal organization and Neurexin and Neuroligin to active zone apposition (Fig. 3j).

In the accompanying manuscript⁸, we showed that although the basal Teneurins are broadly expressed in the *Drosophila* antennal lobe, elevated expression in select glomeruli mediates olfactory neuron partner matching. At the NMJ, this basal level mediates synapse organization. Analogous to the antennal lobe, we found elevated *ten-m* expression at muscles 3 and 8 using the *ten-m-GAL4* enhancer trap (Fig. 4a). We confirmed this for endogenous *ten-m*, and determined that it was contributed by elevated Ten-m expression in both nerves and muscles (Fig. 4b–g). Indeed, *ten-m-GAL4* was highly expressed in select motor neurons, including MN3-Ib, which innervates muscle 3 (ref. 29; Supplementary Fig. 11c). This elevated larval expression also varied along the anterior–posterior axis (Supplementary Fig. 12), and was specific for Ten-m, as Ten-a expression did not differ within or between segments (data not shown).

To test whether elevated Ten-m expression in muscle 3 and MN3-Ib affects neuromuscular connectivity, we expressed *ten-m* RNAi using *ten-m-GAL4*. Wild-type muscle 3 was almost always innervated (Fig. 4h). However, after *ten-m* knockdown, muscle 3 innervation failed in 11% of hemisegments (Fig. 4i, j). This required Ten-m on both sides of the synapse, as the targeting phenotype persisted following neuronal or muscle RNAi suppression using tissue-specific *GAL80* transgenes (Fig. 4j). *ten-a* RNAi did not show this phenotype (Fig. 4j), consistent with homophilic target selection via Ten-m. The phenotype



was specific to muscle 3, as innervation onto the immediately proximal or distal muscle was unchanged (Fig. 4j). The low penetrance is probably due to redundant target selection mechanisms³⁰. Where innervation did occur, the terminal displayed similarly severe phenotypes to other NMJs (not shown). Thus, in addition to generally mediating synaptic organization, Ten-m also contributes to correct target selection at a specific NMJ.

To determine whether Ten-m overexpression could alter connectivity, we expressed Ten-m in muscle 6 (but not the adjacent muscle 7), and the motor neurons innervating both muscles using *H94-GAL4*. Normally, 60% of the boutons at muscles 6/7 are present on muscle 6 with 40% on muscle 7 (Fig. 4k, m). Ten-m overexpression caused a shift whereby 81% of boutons synapsed onto muscle 6 and only 19% onto muscle 7 (Fig. 4l, m). This shift also required both neuronal and muscle Ten-m, as neuronal or muscle *GAL80* abrogated it (Fig. 4m). The effect was specific because Ten-a overexpression did not alter this synaptic balance (Fig. 4m), nor was it secondary to altered bouton

number, which was unchanged (data not shown). Therefore, elevated Ten-m on both sides of the NMJ can bias target choice. This, combined with evidence that Ten-m can mediate homophilic interaction *in vitro*⁸, supports a trans-synaptic homophilic attraction model at the NMJ as in the olfactory system.

We identified a two-tier mechanism for Teneurin function in synapse development at the *Drosophila* NMJ. At the basal level, Teneurins are expressed at all synapses and engage in hetero- and homophilic bi-directional trans-synaptic signalling to organize synapses properly (Fig. 3j). Supporting this, Teneurins can mediate homo- and heterophilic interactions *in vitro*⁸ and heterophilic interactions *in vivo* (Fig. 1f). At the synapse, Teneurins organize the cytoskeleton, interact with α -Spectrin, and enable proper adhesion and release site formation. Further, elevated Ten-m expression regulates target selection in specific motor neurons and muscles via homophilic matching and functions with additional molecules³⁰ to mediate precise neuromuscular connectivity. Teneurin-mediated target selection at the NMJ is analogous to its role in olfactory synaptic partner matching⁸. As Teneurins are expressed broadly throughout the antennal lobe, it remains an attractive possibility that they also regulate synapse organization in the central nervous system.

METHODS SUMMARY

Details of *Drosophila* stocks, immunostaining, electron microscopy, functional assays, construction of epitope-tagged Teneurin constructs, immunoprecipitation, imaging and statistical analysis can be found in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 June 2011; accepted 7 February 2012.

Published online 18 March 2012.

- Williams, M. E., de Wit, J. & Ghosh, A. Molecular mechanisms of synaptic specificity in developing neural circuits. *Neuron* **68**, 9–18 (2010).
- Giagtzoglou, N., Ly, C. V. & Bellen, H. J. Cell adhesion, the backbone of the synapse: “vertebrate” and “invertebrate” perspectives. *Cold Spring Harb. Perspect. Biol.* **1**, a003079 (2009).
- Young, T. R. & Leamey, C. A. Teneurins: important regulators of neural circuitry. *Int. J. Biochem. Cell Biol.* **41**, 990–993 (2009).
- Kenzelmann, D., Chiquet-Ehrismann, R., Leachman, N. T. & Tucker, R. P. Teneurin-1 is expressed in interconnected regions of the developing brain and is processed *in vivo*. *BMC Dev. Biol.* **8**, 30 (2008).
- Li, H., Bishop, K. M. & O’Leary, D. D. Potential target genes of EMX2 include *Odz/Ten-M* and other gene families with implications for cortical patterning. *Mol. Cell. Neurosci.* **33**, 136–149 (2006).
- Silva, J. P. et al. Latrophilin 1 and its endogenous ligand Lasso/teneurin-2 form a high-affinity transsynaptic receptor pair with signaling capabilities. *Proc. Natl Acad. Sci. USA* **108**, 12113–12118 (2011).
- Leamey, C. A. et al. Ten-m3 regulates eye-specific patterning in the mammalian visual pathway and is required for binocular vision. *PLoS Biol.* **5**, e241 (2007) CrossRef.
- Hong, W., Mosca, T. J. & Luo, L. Teneurins instruct synaptic partner matching in an olfactory map. *Nature* <http://dx.doi.org/10.1038/nature10926> (this issue).
- Liebl, F. L. et al. Genome-wide P-element screen for *Drosophila* synaptogenesis mutants. *J. Neurobiol.* **66**, 332–347 (2006).
- Kurusu, M. et al. A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection. *Neuron* **59**, 972–985 (2008).
- Zheng, L. et al. *Drosophila* Ten-m and Filamin affect motor neuron growth cone guidance. *PLoS ONE* **6**, e22956 (2011).
- Sone, M. et al. Synaptic development is controlled in the periaxial zones of *Drosophila* synapses. *Development* **127**, 4157–4168 (2000).
- Wagh, D. A. et al. Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in *Drosophila*. *Neuron* **49**, 833–844 (2006).
- Potter, C. J., Tasic, B., Russler, E. V., Liang, L. & Luo, L. The Q system: a repressible binary system for transgene expression, lineage tracing, and mosaic analysis. *Cell* **141**, 536–548 (2010).
- Marrus, S. B., Portman, S. L., Allen, M. J., Moffat, K. G. & DiAntonio, A. Differential localization of glutamate receptor subunits at the *Drosophila* neuromuscular junction. *J. Neurosci.* **24**, 1406–1415 (2004).
- Wichmann, C. & Sigrist, S. J. The active zone T-bar—a plasticity module? *J. Neurogenet.* **24**, 133–145 (2010).
- Aberle, H. et al. *wishful thinking* encodes a BMP type II receptor that regulates synaptic growth in *Drosophila*. *Neuron* **33**, 545–558 (2002).
- Owald, D. et al. A Syd-1 homologue regulates pre- and postsynaptic maturation in *Drosophila*. *J. Cell Biol.* **188**, 565–579 (2010).
- Kittel, R. J. et al. Bruchpilot promotes active zone assembly, Ca²⁺ channel clustering, and vesicle release. *Science* **312**, 1051–1054 (2006).
- Roos, J., Hummel, T., Ng, N., Klambt, C. & Davis, G. W. *Drosophila* Futsch regulates synaptic microtubule organization and is necessary for synaptic growth. *Neuron* **26**, 371–382 (2000).
- Pielage, J., Fetter, R. D. & Davis, G. W. A postsynaptic spectrin scaffold defines active zone size, spacing, and efficacy at the *Drosophila* neuromuscular junction. *J. Cell Biol.* **175**, 491–503 (2006).
- Pielage, J., Bulat, V., Zuchero, J. B., Fetter, R. D. & Davis, G. W. Hts/Adducin controls synaptic elaboration and elimination. *Neuron* **69**, 1114–1131 (2011).
- Mosca, T. J. & Schwarz, T. L. The nuclear import of Frizzled2-C by Importins- β 1 and α 2 promotes postsynaptic development. *Nature Neurosci.* **13**, 935–943 (2010).
- Featherstone, D. E., Davis, W. S., Dubreuil, R. R. & Broadie, K. *Drosophila* α - and β -spectrin mutations disrupt presynaptic neurotransmitter release. *J. Neurosci.* **21**, 4215–4224 (2001).
- Miech, C., Pauer, H. U., He, X. & Schwarz, T. L. Presynaptic local signaling by a canonical wingless pathway regulates development of the *Drosophila* neuromuscular junction. *J. Neurosci.* **28**, 10875–10884 (2008).
- Craig, A. M. & Kang, Y. Neurexin-neuroligin signaling in synapse development. *Curr. Opin. Neurobiol.* **17**, 43–52 (2007).
- Li, J., Ashley, J., Budnik, V. & Bhat, M. A. Crucial role of *Drosophila* Neurexin in proper active zone apposition to postsynaptic densities, synaptic growth, and synaptic transmission. *Neuron* **55**, 741–755 (2007).
- Banovic, D. et al. *Drosophila* Neuroligin 1 promotes growth and postsynaptic differentiation at glutamatergic neuromuscular junctions. *Neuron* **66**, 724–738 (2010).
- Landgraf, M., Roy, S., Prokop, A., VijayRaghavan, K. & Bate, M. *even-skipped* determines the dorsal growth of motor axons in *Drosophila*. *Neuron* **22**, 43–52 (1999).
- Sanes, J. R. & Yamagata, M. Many paths to synaptic specificity. *Annu. Rev. Cell Dev. Biol.* **25**, 161–195 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. Aberle, V. Budnik, A. DiAntonio, R. Dubreuil, D. Featherstone, N. Reist, T. Schwarz, S. Stowers, D. Van Vactor, R. Wides, the Bloomington Stock Center and the Developmental Studies Hybridoma Bank for fly stocks, antibodies and reagents; J. Perrino and D. Luginbuhl for technical assistance; K. Shen, K. Zinn, D. Banovic, D. Berns, Y. Chou, C. A. Frank, X. Gao, S. Hippenmeyer, K. Miyamichi, K. Sillar, B. Tasic, X. Yu and S. Zosimus for critiques. Supported by a National Institutes of Health (NIH) grant (R01 DC-005982 to L.L.), and Epilepsy, Neonatology and Developmental Biology Training Grants (NIH 5T32 NS007280 and HD007249 to T.J.M.). L.L. is an investigator of the Howard Hughes Medical Institute.

Author Contributions T.J.M. designed and performed all experiments (apart from electrophysiology). W.H. characterized and provided new reagents, and assisted in some experiments. V.S.D. and T.J.M. designed and V.S.D. performed electrophysiology experiments with assistance from T.J.M. V.F. provided new reagents. L.L. supervised the project. T.J.M. wrote the manuscript with feedback from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.J.M. (tmosca@stanford.edu).

METHODS

Drosophila stocks. All *Drosophila* strains and controls were raised at 29 °C to maximize GAL4 expression. All mutants and transgenes were maintained over GFP balancer chromosomes to enable larval selection. *Mhc-GAL4* or *Mef2-GAL4* (ref. 31) was used to drive expression in all somatic muscles. *Nrv2-GAL4* (ref. 32) and *elav-GAL4* (ref. 33) were used to drive expression in all neurons. *H94-GAL4* was used to drive expression in muscles 6, 13 and 4 and their corresponding motor neurons³⁴. *daughterless-GAL4* was used to drive expression ubiquitously³⁵. *Synj-QF*³⁶ was used to drive expression in all nerves. *NP6658-GAL4* (*ten-m-GAL4*) was used to drive expression in the pattern of endogenous *Ten-m* expression⁸. The *Df(X)ten-a* deletion was used as a *ten-a* null mutant⁸. For *nlg1* mutants, the 1960 and ex2.3 alleles were used in *trans*²⁸ and double mutant larvae with *ten-a* and *nlg1* mutations were obtained using optimized rearing conditions³⁷. Because of the early lethality of the *ten-m* mutant¹¹, and to assess independently *ten-a*, tissue-specific RNAi was used to examine *teneurin* perturbation using the following RNAi transgenic strains: for *ten-m*, *UAS-ten-m*^{RNAi-V51173} and for *ten-a* (ref. 8), *UAS-ten-a*^{RNAi-V32482}. The following transgenic strains were used: *UAS-Dcr2* (ref. 38), *UAS-Fas2* (ref. 34), *UAS-mCD8GFP* (ref. 39), *UAS-Nlg1-eGFP* (ref. 28), *UAS-Ten-a* (ref. 8), *P[GS]9267* for *ten-m* overexpression⁸. In all cases, the efficacy of RNAi transgenes, overexpression transgenes and the *ten-a* deletion mutant were assessed and verified by the alteration of antibody staining (loss, reduction or increase) using tissue-specific *GAL4* drivers. For all cases, N-IR indicates neuronal RNAi, M-IR indicates muscle RNAi, U-IR indicates ubiquitous RNAi. For rescuing *ten-a* mutants, Ten-a N indicates neuronal overexpression with *elav-GAL4* of *UAS-ten-a* and Ten-a M indicates muscle overexpression of *UAS-ten-a* with *Mef2-GAL4*.

Immunostaining. Wandering third instar larvae were processed as previously described²³. The following primary antibodies were used: mouse antibody to *Ten-m* (mAb20, 1:500)⁴⁰, guinea pig antibody to *Ten-a* (1:100)⁴¹, mouse antibody to Brp (mAbnc82, 1:250)⁴², rabbit antibody to Synaptotagmin 1 (1:4,000)⁴³, mouse antibody to Cysteine String Protein (mAb6D6, 1:100)⁴⁴, mouse antibody to Dlg (mAb4F3, 1:500)⁴⁵, rabbit antibody to Dlg (1:40,000)⁴⁶, mouse antibody to α -Spectrin (mAb3A9, 1:50)⁴⁷, mouse antibody to Fasciclin 2 (mAb1D4, 1:20)⁴⁸, rabbit antibody to Fasciclin 2 (1:5,000)⁴⁶, mouse antibody to Futsch (mAb22C10, 1:50)²⁰, rabbit antibody to DGluRIII (1:2,500)¹⁵, rat antibody to *Elav* (mAb7E8A10, 1:25)⁴⁹, mouse antibody to Even-skipped (mAb3C10, 1:100)⁵⁰, rabbit antibody to β -Spectrin (1:1,000)⁵¹, mouse antibody to Hts (1:500)²², guinea pig antibody to *Wsp* (1:1,000)⁵². Alexa488-, Alexa546- or Alexa647-conjugated secondary antibodies were used at 1:250 (Invitrogen). Texas-Red-conjugated Phalloidin was used at 1:300. FITC-, Cy3- or Cy5-conjugated antibodies to HRP were used at 1:100 (Jackson ImmunoResearch).

Electron microscopy. Wandering third instar larvae were processed and sectioned as described²³. Sections were imaged on a JEM-1400 (JEOL) transmission electron microscope at $\times 3,000$ to $\times 20,000$ magnification.

Electrophysiology. Larvae were dissected in HL3 saline⁵³ containing 0 mM Ca^{2+} and 4 mM Mg^{2+} . They were then transferred to saline containing 0.6 mM Ca^{2+} and recordings conducted by impaling larval muscle 6 in body wall segments A3 and A4 using sharp intracellular electrodes (10–20 M Ω), fabricated from borosilicate glass capillaries and filled with 3 M KCl solution. For evoked EPSPs, severed nerve bundles were stimulated using a suction electrode connected to a linear stimulus isolated (A395, World Precision Instruments). Data, acquired using Multiclamp 700B amplifiers (Molecular Devices), were low-pass filtered at 3 kHz and digitized at 10 kHz. Recordings were acquired and analysed using Igor Pro software (Wavemetrics) and custom-written programs. All recordings in which the resting membrane potential was higher than -60 mV and/or whose resting potential, input resistance or access resistance changed by more than 20% during the duration of data acquisition were excluded from analysis. All recordings and data analyses were performed blind to the genotype. Quantal content was corrected for nonlinear summation⁵⁴.

Larval locomotion. Crawling assays were conducted as described⁵⁵.

FM1-43 dye loading. FM1-43 (Invitrogen) dye loading was conducted as described⁵⁶ with the following modifications: loading was conducted in 1.5 mM Ca^{2+} , 90 mM K^{+} saline for 1 min followed by six 2-min washes in 0 mM Ca^{2+} saline. Imaging was conducted on a Zeiss LSM 510 Meta Confocal (Carl Zeiss) with a $\times 40$, PlanApo NA 1.0 water immersion lens (Carl Zeiss).

Construction of epitope-tagged teneurin transgenes. The *ten-m* and *ten-a* coding sequences lacking the stop codon were cloned into *pENTR-D/TOPO* (Invitrogen) from *pENTR-ten-m* and *pENTR-ten-a*. *pENTR-ten-m* (stop) and *pENTR-ten-a* (stop) were recombined into the destination vectors *pUASTattB-gtw-tFHAH* and *pUASTattB-gtw-tFMH*, respectively, using LR Clonase II (Invitrogen). *pUASTattB-gtw-tFHAH* is a *pUAST-Gateway-attB* based vector with a C-terminal TEV recognition site and 3 \times Flag, 3 \times HA and 10 \times His tags. *pQUASTattB-gtw-tFMH* is a *pQUAST-Gateway-attB*⁸ based vector with a

C-terminal TEV recognition site and 3 \times Flag, 6 \times Myc and 10 \times His tags. The resulting constructs were verified by restriction digest and sequencing and integrated into the *attP24* or *86Fb* landing sites on the second and third chromosomes⁵⁷. Transgenic flies were verified by immunoprecipitation on western blot and overexpression experiments.

Immunoprecipitation, western blots and SDS-PAGE analysis. For *Ten-m* and *Ten-a*, *QUAS-Ten-a-Flag-Myc* was expressed in nerves using *Synj-QF* and *UAS-Ten-m-Flag-HA* in muscles using *mhc-GAL4*. Larval synaptosomes were prepared from larval body wall fillets as described⁵⁸. For *Ten-m* and α -Spectrin, control larvae consisted of *Mef2-GAL4* without *UAS-Ten-m-Flag-HA* whereas experimental flies combined the two. Immunoprecipitation was conducted as described using M2-anti-Flag-conjugated agarose (Sigma) or Affi-Prep Protein A beads (Bio-Rad) and rat antibodies to HA (Roche)²³. Proteins were separated on NuPAGE 3–8% Tris-Acetate Gels (Invitrogen) and transferred to nitrocellulose. Primary antibodies were applied overnight at 4 °C and secondary antibodies at 21 °C for 1 h. The following primary antibodies were used: mouse antibody to α -Spectrin (mAb3A9, 1:2,000), mouse antibody to Brp (mAbnc82, 1:100), mouse antibody to Flag (M2, 1:5,000, Sigma-Aldrich), mouse antibody to Myc (3E10, 1:1,500, Santa Cruz Biotechnology), rat antibody to HA (3F10, 1:1,500, Roche). HRP-conjugated secondary antibodies (Jackson ImmunoResearch) were used at 1:10,000. Blots were developed using the SuperSignal West Femto Maximum Sensitivity Substrate (ThermoScientific).

Imaging analysis. Larvae were imaged with a Zeiss LSM 510 Meta laser-scanning confocal microscope (Carl Zeiss) using either a $\times 63$ 1.4 NA or a $\times 40$ 1.0 NA objective. NMJ images were taken as confocal z-stacks with the upper and lower bounds defined by HRP staining unless otherwise noted. For all metrics, boutons were assessed in segment A3 at muscle 6/7 and muscle 4 on both the left and right sides. Fluorescence intensity measurements were taken from terminals on muscle 4. All phenotypes, however, were observed at all synapses regardless of muscle fibre or segment. For membrane organization, vesicle distribution and Teneurin colocalization, NMJ images were taken as single optical sections at the precise centre of the bouton as determined by HRP staining. Images were processed with the LSM software and Adobe Photoshop CS4. Bouton number, active zone/glutamate receptor apposition, fluorescent intensity and microtubule organization were quantified as previously described²³. Targeting errors for each larva were quantified as the percentage of hemisegments from A1 to A7 in a single animal with a failure of target innervation. There was no difference in targeting errors based on body wall segment. Experiments using *H94-GAL4* were conducted as described³⁴, and their effects confirmed using Fasciclin 2 overexpression (control = 58.1% of boutons on muscle 6, 41.9% on muscle 7; Fas 2 overexpression = 73.0% on muscle 6, 27.0% on muscle 7; $n = 8$ animals for each, $P < 0.0001$)³⁴.

In electron micrographs, parameters were quantified as previously described using ImageJ (NIH)²³. T-bar defects were classified into one of five categories: normal (no discernible defect), double (two T-bars were observed in the same, continuous active zone), detached (where the T-bar was clearly visible but was not explicitly connected to the membrane associated with the nearest PSD), apposite contractile tissue (where the T-bar was not apposed to the SSR, but rather, the contractile tissue of the muscle), misshapen (where an electron-dense T-bar was visible but did not conform to the 'T' shape. Often, the T-bars were 'X' shaped). For Fig. 2r, each defect is expressed as a percentage of the total number of T-bars observed in a particular genotype.

For *Ten-m* gradient calculation, single optical sections were taken through the centre of the NMJ on muscle 3 or muscle 4, as determined by HRP immunoreactivity. The GFP signal (*ten-m-GAL4*) or antibody signal was then measured using ImageJ (NIH). For each larva, measurements were taken on the right and left sides of each indicated segment. The fluorescence for each segment was expressed as a percentage of the fluorescence from segment A1 in the same animal, on the same side of the larvae. For all larvae, segment A1 represented the maximal fluorescence. **Statistical analysis.** Statistical analysis used GraphPad Prism 5 (Graphpad Software). In all cases involving more than two samples, significance was calculated using ANOVA followed by a Dunnett post-hoc test to the control sample and a Bonferroni post-hoc test among all samples. For two-sample cases, an unpaired Student's *t*-test was used to assess significance, unless otherwise indicated. In all cases, both methods provided similar significance measurements. In all figures, significance is with respect to control genotypes unless otherwise noted.

- Lilly, B. et al. Requirement of MADS domain transcription factor D-MEF2 for muscle formation in *Drosophila*. *Science* **267**, 688–693 (1995).
- Sun, B., Xu, P. & Salvaterra, P. M. Dynamic visualization of nervous system in live *Drosophila*. *Proc. Natl Acad. Sci. USA* **96**, 10438–10443 (1999).
- Luo, L., Liao, Y. J., Jan, L. Y. & Jan, Y. N. Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes Dev.* **8**, 1787–1802 (1994).
- Davis, G. W. & Goodman, C. S. Synapse-specific control of synaptic efficacy at the terminals of a single neuron. *Nature* **392**, 82–86 (1998).

35. Wodarz, A., Hinz, U., Engelbert, M. & Knust, E. Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of *Drosophila*. *Cell* **82**, 67–76 (1995).
36. Petersen, L. K. & Stowers, R. S. A Gateway MultiSite recombination cloning toolkit. *PLoS ONE* **6**, e24531 (2011).
37. Loewen, C. A., Mackler, J. M. & Reist, N. E. *Drosophila* *synaptotagmin I* null mutants survive to early adulthood. *Genesis* **31**, 30–36 (2001).
38. Dietzl, G. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151–156 (2007).
39. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
40. Levine, A. *et al.* *odd Oz*: a novel *Drosophila* pair rule gene. *Cell* **77**, 587–598 (1994).
41. Rakovitsky, N. *et al.* *Drosophila* *Ten-a* is a maternal pair-rule and patterning gene. *Mech. Dev.* **124**, 911–924 (2007).
42. Laissue, P. P. *et al.* Three-dimensional reconstruction of the antennal lobe in *Drosophila melanogaster*. *J. Comp. Neurol.* **405**, 543–552 (1999).
43. Mackler, J. M., Drummond, J. A., Loewen, C. A., Robinson, I. M. & Reist, N. E. The Ca^{2+} -binding motif of synaptotagmin is required for synaptic transmission *in vivo*. *Nature* **418**, 340–344 (2002).
44. Zinsmaier, K. E., Eberle, K. K., Buchner, E., Walter, N. & Benzer, S. Paralysis and early death in cysteine string protein mutants of *Drosophila*. *Science* **263**, 977–980 (1994).
45. Parnas, D., Haghighi, A. P., Fetter, R. D., Kim, S. W. & Goodman, C. S. Regulation of postsynaptic structure and protein localization by the Rho-type guanine nucleotide exchange factor dPix. *Neuron* **32**, 415–424 (2001).
46. Koh, Y. H., Popova, E., Thomas, U., Griffith, L. C. & Budnik, V. Regulation of DLG localization at synapses by CaMKII-dependent phosphorylation. *Cell* **98**, 353–363 (1999).
47. Byers, T. J., Dubreuil, R., Branton, D., Kiehart, D. P. & Goldstein, L. S. *Drosophila* spectrin. II. Conserved features of the alpha-subunit are revealed by analysis of cDNA clones and fusion proteins. *J. Cell Biol.* **105**, 2103–2110 (1987).
48. Van Vactor, D., Sink, H., Fambrough, D., Tsao, R. & Goodman, C. S. Genes that control neuromuscular specificity in *Drosophila*. *Cell* **73**, 1137–1153 (1993).
49. O'Neill, E. M., Rebay, I., Tjian, R. & Rubin, G. M. The activities of two Ets-related transcription factors required for *Drosophila* eye development are modulated by the Ras/MAPK pathway. *Cell* **78**, 137–147 (1994).
50. Patel, N. H., Schafer, B., Goodman, C. S. & Holmgren, R. The role of segment polarity genes during *Drosophila* neurogenesis. *Genes Dev.* **3**, 890–904 (1989).
51. Byers, T. J., Husain-Chishti, A., Dubreuil, R. R., Branton, D. & Goldstein, L. S. Sequence similarity of the amino-terminal domain of *Drosophila* beta spectrin to alpha actinin and dystrophin. *J. Cell Biol.* **109**, 1633–1641 (1989).
52. Ben-Yaacov, S., Le Borgne, R., Abramson, I., Schweisguth, F. & Schejter, E. D. *Wasp*, the *Drosophila* Wiskott-Aldrich syndrome gene homologue, is required for cell fate decisions mediated by *Notch* signaling. *J. Cell Biol.* **152**, 1–13 (2001).
53. Stewart, B. A., Atwood, H. L., Renger, J. J., Wang, J. & Wu, C. F. Improved stability of *Drosophila* larval neuromuscular preparations in haemolymph-like physiological solutions. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **175**, 179–191 (1994).
54. Martin, A. R. A further study of the statistical composition on the end-plate potential. *J. Physiol. (Lond.)* **130**, 114–122 (1955).
55. Lnenicka, G. A., Spencer, G. M. & Keshishian, H. Effect of reduced impulse activity on the development of identified motor terminals in *Drosophila* larvae. *J. Neurobiol.* **54**, 337–345 (2003).
56. Verstreken, P., Ohyama, T. & Bellen, H. J. FM 1-43 labeling of synaptic vesicle pools at the *Drosophila* neuromuscular junction. *Methods Mol. Biol.* **440**, 349–369 (2008).
57. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genet.* **40**, 476–483 (2008).
58. Higashi-Kovtun, M. E., Mosca, T. J., Dickman, D. K., Meinertzhagen, I. A. & Schwarz, T. L. Importin- β 11 regulates synaptic phosphorylated mothers against decapentaplegic, and thereby influences synaptic development and function at the *Drosophila* neuromuscular junction. *J. Neurosci.* **30**, 5253–5268 (2010).

Transcription factor PIF4 controls the thermosensory activation of flowering

S. Vinod Kumar^{1*}, Doris Lucyshyn^{1†*}, Katja E. Jaeger¹, Enriqueta Alós¹, Elizabeth Alvey¹, Nicholas P. Harberd^{1,2} & Philip A. Wigge^{1†}

Plant growth and development are strongly affected by small differences in temperature¹. Current climate change has already altered global plant phenology and distribution^{2,3}, and projected increases in temperature pose a significant challenge to agriculture⁴. Despite the important role of temperature on plant development, the underlying pathways are unknown. It has previously been shown that thermal acceleration of flowering is dependent on the florigen, FLOWERING LOCUS T (FT)^{5,6}. How this occurs is, however, not understood, because the major pathway known to upregulate FT, the photoperiod pathway, is not required for thermal acceleration of flowering⁶. Here we demonstrate a direct mechanism by which increasing temperature causes the bHLH transcription factor PHYTOCHROME INTERACTING FACTOR4 (PIF4) to activate FT. Our findings provide a new understanding of how plants control their timing of reproduction in response to temperature. Flowering time is an important trait in crops as well as affecting the life cycles of pollinator species. A molecular understanding of how temperature affects flowering will be important for mitigating the effects of climate change.

Arabidopsis thaliana, like many higher plants, responds to warmer ambient temperatures by increasing its growth rate and accelerating the floral transition^{1,5,7}. *Arabidopsis* is a facultative long-day plant, and plants grown under short photoperiods are dramatically delayed in flowering. Interestingly, late flowering in short days can be overcome by growth at higher temperatures⁶. The underlying mechanism is, however, unknown. The flowering response to temperature is dependent on the floral pathway integrator gene *FT*⁶, indicative of a thermosensory pathway that upregulates FT expression independently of daylength. Because the bHLH transcription factor PIF4 has been shown to regulate architectural responses to high temperature^{8,9}, we tested if *PIF4* is required for the induction of flowering at high temperature in short photoperiods. Although *pif4-101* was slightly delayed in flowering at 22 °C, *pif4-101* mutants showed a striking loss of thermal induction of flowering at 27 °C (Fig. 1a, b). To test if *pif4-101* perturbed floral induction by affecting FT expression, we examined the thermal induction of FT in Col-0 and *pif4-101*. Although FT expression was strongly thermally inducible in Col-0, this response was largely abolished in *pif4-101* at 27 °C (Fig. 1c), indicating that *PIF4* is necessary for the thermal acceleration of flowering in short days. By contrast, *PIF4* is not required for the thermosensory induction of flowering under continuous light⁸, suggesting that the photoperiod pathway also interacts with the ambient temperature sensing pathway. The reduced role of *PIF4* under continuous light probably reflects the instability of PIF4 in light¹⁰ coupled with the fact that the output of the photoperiod pathway, CONSTANS (CO) protein, is stabilized by light¹¹, shifting the balance of floral induction from *PIF4* to the photoperiod pathway. Because *PIF4* is necessary for the thermal induction of flowering in short days, we tested if it is sufficient to trigger flowering when overexpressed. *35S::PIF4* caused extremely early flowering

(Fig. 1d, e), similar to the effect of overexpressing a related gene, PHYTOCHROME INTERACTING FACTOR5 (ref. 12), suggesting that *PIF4* is limiting for the acceleration of flowering at lower temperature in short photoperiods. Consistently, *35S::PIF4* plants showed elevated levels of FT (Fig. 1f). Furthermore, *35S::PIF4 ft-3* showed a complete suppression of the early flowering phenotype, indicating that the induction of flowering by *35S::PIF4* was dependent on FT (Fig. 1g, h). This activation of FT appears to be independent of the established

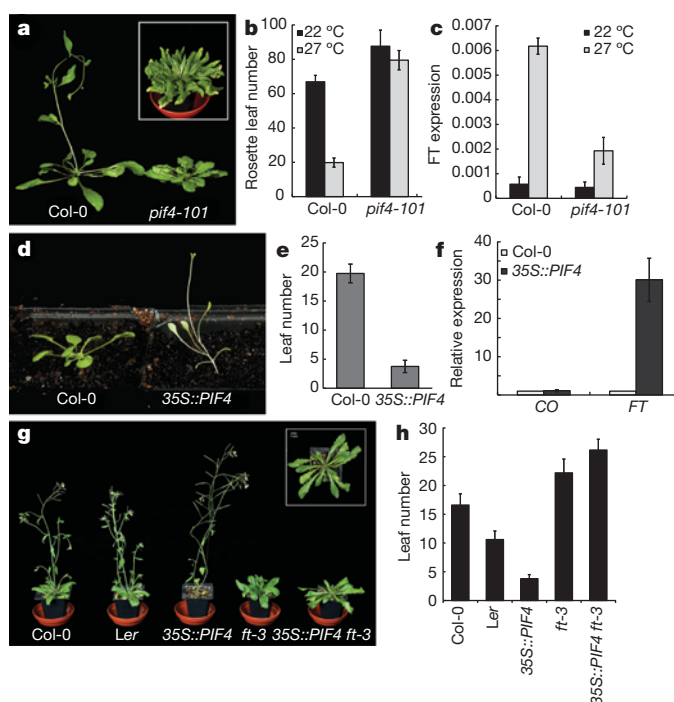


Figure 1 | *PIF4* is necessary for the thermal induction of flowering in short photoperiods. **a**, *pif4-101* plants do not show acceleration of flowering at 27 °C compared with Col-0. Inset, a 16-week-old *pif4-101* plant grown at 27 °C. **b**, Rosette leaf numbers at flowering for Col-0 and *pif4-101* grown at 22 and 27 °C in short-photoperiod conditions (error bars, SD; $n = 6$). **c**, FT expression as measured by quantitative polymerase chain reaction (qPCR) in 4-week-old plants at 22 and 27 °C under short photoperiods in a *PIF4*-dependent manner (data from three biological replicates; error bars, SD). **d**, *35S::PIF4* overexpression triggers very early flowering. **e**, Leaf numbers at flowering for Col-0 and *35S::PIF4* in long photoperiods (error bars, SD; $n = 5$). **f**, CO and FT gene expression data measured by qPCR in Col-0 and *35S::PIF4* at 21 °C in long photoperiods (samples taken 2 weeks after sowing; data from three biological replicates; all error bars, SD). **g**, FT is required for the early flowering phenotype of *35S::PIF4* plants. When crossed into the *ft-3* background, the early flowering of *35S::PIF4* is completely suppressed. Inset, top view of the *35S::PIF4 ft-3* plant showing that petiole elongation growth is retained. **h**, Flowering time data for Col-0, *Ler*, *35S::PIF4*, *ft-3* and *35S::PIF4 ft-3* plants (error bars, SD; $n = 5$).

¹Department of Cell and Developmental Biology, John Innes Centre, Norwich NR4 7UH, UK. ²Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK. [†]Present addresses: Institute for Applied Genetics and Cell Biology, BOKU—University for Applied Life Sciences and Natural Resources, Muthgasse 18, 41190 Vienna, Austria (D.L.); Sainsbury Laboratory, Cambridge University, Bateman Street, Cambridge CB2 1LR, UK (P.A.W.).

*These authors contributed equally to this work.

photoperiod pathway because *CO* did not change in response to *35S::PIF4* (Fig. 1f). Finally, although *co-9* mutants are late flowering^{6,13}, we found *35S::PIF4 co-9* plants were early flowering, indicating that *PIF4* acts largely independently of *CO* (Supplementary Information and Supplementary Fig. 1), consistent with the thermal induction of flowering being independent of the photoperiod pathway (Fig. 1f)⁶.

Although *PIF4* has been shown to be important for high-temperature responses, long-term increases in either *PIF4* transcript or *PIF4* protein levels in response to higher ambient temperature that can account for the observed growth responses have not been detected^{8,9}. To examine if variation of *PIF4* transcription under our experimental conditions might account for the increases in *PIF4* activity with temperature, we measured *PIF4* transcript levels at 12, 17, 22 and 27 °C in seedlings (Fig. 2a). *PIF4* transcript levels increased from 12 °C to 22 °C, whereas the difference between 22 °C and 27 °C was not statistically significant. Plants at 27 °C, compared with 22 °C, showed a very large *PIF4*-dependent response, suggesting that variation in the *PIF4* transcript is not sufficient to account for the acceleration of flowering at 27 °C compared with 22 °C. To test whether temperature-mediated changes in *PIF4* transcription are rate limiting for the biological response, we analysed the behaviour of plants constitutively expressing *PIF4*. Although *35S::PIF4* plants at 22 °C were extremely early flowering, this phenotype could be largely suppressed at 12 °C (Fig. 2b and Supplementary Fig. 2), indicating that even when *PIF4* transcript is abundant, lower temperatures are inhibitory for *PIF4* activity. A possible explanation for this difference is that *PIF4* protein is destabilized by low temperature. Indeed, *PIF4* protein levels have already been shown to be strongly regulated by light¹⁰, and growth in red and blue photocycles destabilizes *PIF4* protein at low temperatures¹⁴. The *PIF4* overexpression lines contain a fusion to the haemagglutinin (HA) epitope (*35S::PIF4:HA*). We therefore examined the levels of *PIF4:HA*, protein at 12, 17, 22 and 27 °C under the same light conditions used for our flowering time assays. Consistent with previous studies¹⁰ we saw a strong accumulation of *PIF4* at the end of the night period, which was subsequently degraded during the day. Despite the suppression of early flowering in *35S::PIF4* at 12 °C compared with 22 °C (Fig. 2b), we did not observe an appreciable difference in *PIF4* protein levels at

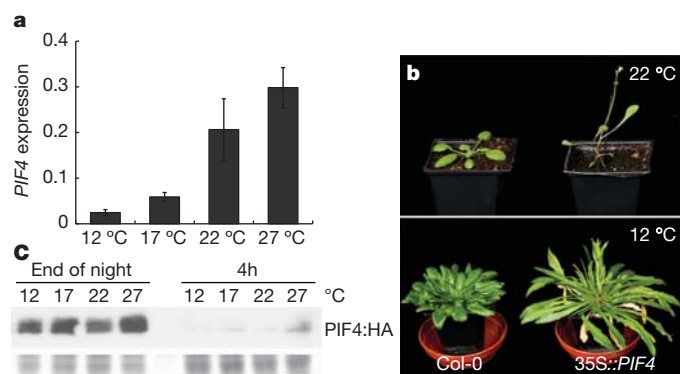


Figure 2 | Regulation of *PIF4* by temperature. **a**, Transcriptional regulation of *PIF4* by temperature. Ten-day-old Col-0 seedlings grown at 12, 17, 22 and 27 °C under short photoperiods were analysed for *PIF4* expression by qPCR. Data shown are from three biological replicates. Error bars, s.d. **b**, Flowering phenotype of *35S::PIF4* is temperature dependent. Although *35S::PIF4* plants (right) flower very early at 22 °C (upper panel) compared with Col-0 (left), this phenotype is largely suppressed by growth at 12 °C (lower panel). **c**, *PIF4* protein levels in *35S::PIF4* plants are not affected by growth temperature. Seven-day-old *35S::PIF4:HA* seedlings grown at 17 °C were transferred to 12, 17, 22 and 27 °C under short photoperiods for 2 days and samples were collected at the end of night before light and after 4 h under illumination. Although *PIF4:HA* protein levels are independent of growth temperatures, the protein is robustly degraded in presence of light. *PIF4:HA* protein was detected by horseradish peroxidase (HRP)-conjugated HA antibody. Stained lower half of the gel used for immunoblot is shown as loading control.

these two temperatures that was likely to account for these different phenotypes (Fig. 2c and Supplementary Fig. 3). Slightly higher levels of *PIF4:HA* appeared to be present at 27 °C (Fig. 2c), suggesting high-temperature stabilization of *PIF4* may also contribute to higher *PIF4* activity at 27 °C.

Taken together, these data indicate that *PIF4* regulates *FT* in a temperature-dependent manner. To determine if this is probably the case *in planta*, we analysed the spatial expression of *FT* and *PIF4*. *FT* has a distinctive pattern of expression in the vasculature of the leaf^{15,16}, and significantly *PIF4* was expressed in the same domain (Fig. 3a). Because the regulation of *FT* by *PIF4* could be either direct or indirect, we used chromatin immunopurification (ChIP) to analyse if *PIF4* binds directly to the *FT* promoter proximal to the transcriptional start site. This region of the promoter was chosen because it has been shown to be both phylogenetically conserved and the site for light-mediated regulation of *FT* expression^{16,17}. We observed robust enrichment of *PIF4* near to the transcriptional start site (Fig. 3b), indicating that *PIF4* binds this region *in vivo* to activate *FT* expression.

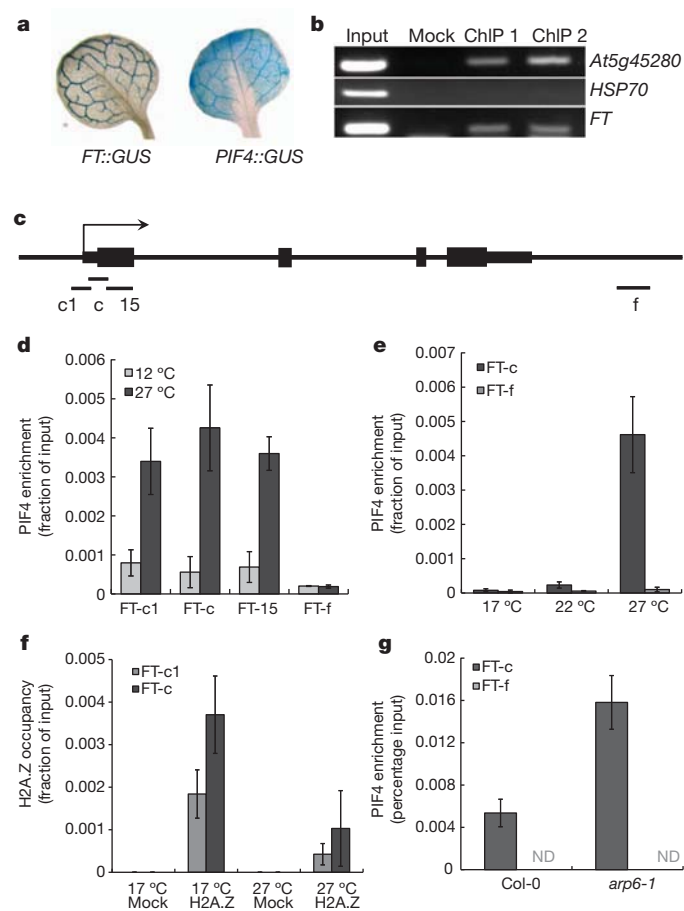


Figure 3 | *PIF4* directly binds the *FT* promoter in a temperature-dependent manner. **a**, β -Glucuronidase (GUS) histochemical analysis of the expression domains of *FT* and *PIF4* in rosette leaves. **b**, ChIP analysis shows *PIF4* binding to the *FT* locus *in vivo* in seedlings. The *At5g45280* promoter is a positive control for *PIF4*-binding activity²¹, the *HSP70* promoter was used as a negative control. **c**, Summary of the structure of the *FT* promoter and positioning of qPCR amplicons for ChIP analysis. **d**, ChIP analysis of *35S::PIF4:HA* at 12 and 27 °C (2-week-old seedlings, short photoperiods). **e**, ChIP analysis of *PIF4::PIF4:ProA* at 17, 22 and 27 °C (4-week-old soil-grown plants, short photoperiods). **f**, Analysis of H2A.Z occupancy at the *FT* locus at 17 and 27 °C (3-week-old plate-grown plants, short photoperiods). **g**, ChIP analysis of *PIF4* binding to *FT* in Col-0 and *arp6-1* (3-week-old soil-grown plants, 22 °C short photoperiod). (For all ChIP experiments, plant materials were collected at the end of dark period before lights came on and were protected from light until frozen. All data presented are from two independent ChIP experiments; all error bars, SD.)

Given the striking effect of ambient temperature on PIF4 activity, which occurs even when *PIF4* is constitutively expressed, we hypothesized that the ability of PIF4 to bind the *FT* promoter may be temperature dependent. To test this, we performed ChIP experiments using 35S::*PIF4* plants grown at 12 and 27 °C with primers flanking an E-box in the *FT* promoter (Fig. 3c). Strikingly, we observed a very strong temperature dependence for this binding, with an approximately fivefold increase in binding at 27 °C compared with 12 °C (Fig. 3d). This indicates that the later flowering of 35S::*PIF4* at 12 °C is caused by a decrease in PIF4 binding to *FT*. Because the 35S promoter causes strong ectopic expression of PIF4, we sought to confirm that PIF4 protein expressed at endogenous levels displays similar temperature-dependent binding to the *FT* promoter. We therefore performed ChIP experiments on a *pif4-101* line complemented with *PIF4_{pro}::PIF4:ProteinA* (Supplementary Fig. 4). Consistent with the overexpression studies, we observed a strong increase in PIF4 binding to *FT* as a function of temperature. Reduced binding was observed at 17 °C, consistent with the very late flowering of plants under short days at low temperature, but this binding increased at 22 °C and was even higher at 27 °C (Fig. 3e). The temperature-dependent binding of PIF4 to *FT* could be due to growth temperature influencing the affinity of the PIF4 transcription factor for its binding site, or the efficiency of the ChIP could be affected by the temperature at which tissues were grown. To test these possibilities, we analysed another recently described PIF4 target locus¹⁸, *CYP79B2* (At4g39950), which is upregulated in 35S::*PIF4* (Supplementary Fig. 5a). We found PIF4 binding to occur constitutively at both 12 and 27 °C at a region in the first exon (Supplementary Fig. 5b). Another region further upstream in the promoter showed a temperature-dependent binding of PIF4, and, in both cases, no enrichment was seen for a control locus (Supplementary Fig. 5b). This indicates that the abundant PIF4 protein we observed at 12 °C is active and able to bind target sites, and confirms that the ChIP method per se is not influenced by the temperature at which the sample is grown, consistent with other studies¹⁹. The ability of PIF4 to bind loci in a more temperature-independent manner might explain why 35S::*PIF4* at 12 °C maintains hypocotyl and petiole elongation, while early flowering is strongly suppressed. We do not exclude that temperature may also influence PIF4 activity post-translationally.

Temperature signals are mediated through H2A.Z-nucleosomes in *Arabidopsis*²⁰, suggesting that temperature may be increasing the accessibility of the PIF4-binding site at the *FT* promoter. Consistent with this hypothesis, we found that H2A.Z-nucleosomes were present at the PIF4-binding site in the *FT* promoter. Furthermore, we found that the levels of H2A.Z-nucleosomes at the *FT* promoter decreased with higher temperature (Fig. 3f). These results suggest that the presence of H2A.Z-nucleosomes is limiting for PIF4 binding to *FT*, and that the PIF4 binding we observed at higher temperature is due to the greater accessibility of chromatin containing H2A.Z-nucleosomes at higher temperature. This suggests that in the absence of H2A.Z-nucleosomes, PIF4 should bind *FT* more strongly. We therefore compared the ability of PIF4 expressed under its own promoter to bind to the *FT* promoter in wild type and *arp6-1*, a background lacking incorporation of H2A.Z-nucleosomes. Interestingly, we observed considerably greater binding of PIF4 in *arp6-1* (Fig. 3g), indicating that H2A.Z-nucleosomes are rate limiting for PIF4 to activate *FT* expression. The eviction of H2A.Z-nucleosomes by higher temperature therefore provides a direct mechanism for the temperature-regulated expression of *FT* (Fig. 4c). Consistent with our previous results and the established role of H2A.Z in regulating temperature-dependent gene expression, we found that there is increased PIF4 messenger RNA in *arp6-1* background (Supplementary Fig. 6). However, our results for 35S::*PIF4* suppression by 12 °C indicate that transcriptional upregulation of PIF4 is not the rate-limiting step in regulating PIF4-mediated flowering at higher temperatures.

Our results indicate that the temperature-dependent regulation of *FT* by PIF4 is controlled at the level of chromatin accessibility of the

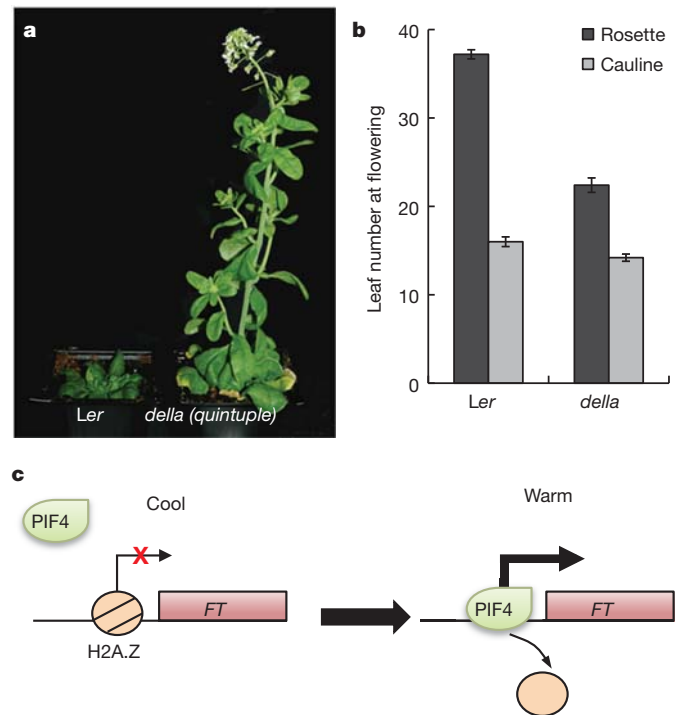


Figure 4 | PIF4 integrates environmental signals. **a**, Suppression of flowering at 12 °C is significantly repressed in the absence of DELLA-mediated repression. **b**, Leaf number at flowering for *della* global is reduced compared with Ler at 12 °C (error bars, SD; $n = 6$). **c**, Representation of temperature-dependent *FT* regulation by PIF4. Temperature-induced H2A.Z nucleosome dynamics can regulate PIF4 binding to target loci for transcriptional activation.

FT promoter and possibly at the level of PIF4 protein activity. PIF4 activity is controlled through the repressive activity of DELLA proteins that prevent PIF4 binding DNA^{21,22}.

Consistently, plants having reduced or absent DELLA function are early flowering²³. We hypothesized that delay in flowering at lower temperatures might at least in part be due to DELLA-mediated repression of PIF4 activity. If so, it would be expected that absence of DELLAs should cause accelerated flowering at lower temperatures. In accord with this expectation, we found that a mutant lacking DELLAs flowered much earlier than wild type when grown at 12 °C (Fig. 4a, b). The phytohormone gibberellin triggers DELLA protein degradation, and plays a key permissive role for *FT* induction, because in a gibberellin-deficient background, gibberellin application increases *FT* expression 15-fold²⁴. Although it was proposed more than 50 years ago that gibberellins are upstream of florigen²⁵, the mechanism has not been clear. As DELLA proteins have been shown to be key regulators by which gibberellin influences PIF4, our finding that PIF4 is able to activate *FT* directly suggests a possible mechanism by which changes in gibberellin levels may influence flowering.

Climate change has already caused measurable changes in plant phenology and behaviour², and plants that incorporate temperature information into their life cycles appear to be able to adapt to warmer conditions more effectively than those that primarily rely on photoperiod to synchronize their lifestyles³. The importance of the effects of climate change on yield are highlighted by the significant detrimental effects of increasing temperatures on yield⁴. *PIF4* is a central integrator of environmental information in the plant and our finding that it activates *FT* at higher temperatures suggests it will be a key node for breeding crops resilient to climate change. This importance is suggested by the recent discovery that natural variation at *PIF4* plays a major role in key ecological traits²⁶.

METHODS SUMMARY

Detailed descriptions of the plant growth conditions, growth assays, transgenic constructs and ChIP techniques are provided in the Supplementary materials.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 July 2010; accepted 6 February 2012.

Published online 21 March 2012.

- Samach, A. & Wigge, P. A. Ambient temperature perception in plants. *Curr. Opin. Plant Biol.* **8**, 483–486 (2005).
- Fitter, A. H. & Fitter, R. S. Rapid changes in flowering time in British plants. *Science* **296**, 1689–1691 (2002).
- Willis, C. G., Ruhfel, B., Primack, R. B., Miller-Rushing, A. J. & Davis, C. C. Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proc. Natl Acad. Sci. USA* **105**, 17029–17033 (2008).
- Battisti, D. S. & Naylor, R. L. Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* **323**, 240–244 (2009).
- Halliday, K. J., Salter, M. G., Thingnaes, E. & Whitelam, G. C. Phytochrome control of flowering is temperature sensitive and correlates with expression of the floral integrator FT. *Plant J.* **33**, 875–885 (2003).
- Balasubramanian, S., Sureshkumar, S., Lempe, J. & Weigel, D. Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. *PLoS Genet.* **2**, e106 (2006).
- Blazquez, M. A., Ahn, J. H. & Weigel, D. A thermosensory pathway controlling flowering time in *Arabidopsis thaliana*. *Nature Genet.* **33**, 168–171 (2003).
- Koini, M. A. *et al.* High temperature-mediated adaptations in plant architecture require the bHLH transcription factor PIF4. *Curr. Biol.* **19**, 408–413 (2009).
- Stavang, J. A. *et al.* Hormonal regulation of temperature-induced growth in *Arabidopsis*. *Plant J.* **60**, 589–601 (2009).
- Nozue, K. *et al.* Rhythmic growth explained by coincidence between internal and external cues. *Nature* **448**, 358–361 (2007).
- Valverde, F. *et al.* Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* **303**, 1003–1006 (2004).
- Lorrain, S., Allen, T., Duek, P. D., Whitelam, G. C. & Fankhauser, C. Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors. *Plant J.* **53**, 312–323 (2008).
- Wenkel, S. *et al.* CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. *Plant Cell* **18**, 2971–2984 (2006).
- Foreman, J. *et al.* Light receptor action is critical for maintaining plant biomass at warm ambient temperatures. *Plant J.* **65**, 441–452 (2011).
- Takada, S. & Goto, K. Terminal flower2, an *Arabidopsis* homolog of heterochromatin protein1, counteracts the activation of flowering locus T by constans in the vascular tissues of leaves to regulate flowering time. *Plant Cell* **15**, 2856–2865 (2003).
- Adrian, J. *et al.* cis-regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in *Arabidopsis*. *Plant Cell* **22**, 1425–1440 (2010).
- Liu, H. *et al.* Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. *Science* **322**, 1535–1539 (2008).
- Franklin, K. A. *et al.* PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) regulates auxin biosynthesis at high temperature. *Proc. Natl Acad. Sci. USA* **108**, 20231–20235 (2011).
- Angel, A., Song, J., Dean, C. & Howard, M. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476**, 105–108 (2011).
- Kumar, S. V. & Wigge, P. A. H2A.Z-containing nucleosomes mediate the thermosensory response in *Arabidopsis*. *Cell* **140**, 136–140 (2010).
- de Lucas, M. *et al.* A molecular framework for light and gibberellin control of cell elongation. *Nature* **451**, 480–484 (2008).
- Feng, S. *et al.* Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins. *Nature* **451**, 475–479 (2008).
- Mutasa-Gottgens, E. & Hedden, P. Gibberellin as a factor in floral regulatory networks. *J. Exp. Bot.* **60**, 1979–1989 (2009).
- Hisamatsu, T. & King, R. W. The nature of floral signals in *Arabidopsis*. II. Roles for FLOWERING LOCUS T (FT) and gibberellin. *J. Exp. Bot.* **59**, 3821–3829 (2008).
- Brian, P. W. Role of gibberellin-like hormones in regulation of plant growth & flowering. *Nature* **181**, 1122–1123 (1958).
- Brock, M. T., Maloof, J. N. & Weinig, C. Genes underlying quantitative variation in ecologically important traits: PIF4 (phytochrome interacting factor 4) is associated with variation in internode length, flowering time, and fruit set in *Arabidopsis thaliana*. *Mol. Ecol.* **19**, 1187–1199 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Prat, C. Fankhauser, K. Franklin, K. Goto, G. Coupland and D. Weigel for seeds. We are grateful to members of the Wigge laboratory for discussions. This work was supported in part by award No. KUK-I1-002-03 (to N.P.H.) made by King Abdullah University of Science and Technology and a Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/I019022/1 (to S.V.K.). D.L. was supported by an Erwin Schroedinger Fellowship from the Austrian Science Fund FWF. P.A.W. was supported by start-up funds from the John Innes Centre and BBSRC, a BBSRC grant (BB/D0100470/1) and a European Research Council Starting Grant (ERC 243140).

Author Contributions S.V.K., D.L., K.E.J. and E.A. performed the experiments. N.P.H. and P.A.W. designed the study and supervised the work. All authors discussed the results and made substantial contributions to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.A.W. (philip.wigge@slcu.cam.ac.uk).

METHODS

Plant material and growth conditions. All plant lines used were in Col-0 background unless otherwise specified. *pif4-101* mutant was provided by C. Fankhauser, HA-tagged 35S::PIF4 by S. Prat¹². All references to '35S::PIF4' and 'PIF4:HA' refer to this line: that is, 35S::PIF4:HA. *FT::GUS* was obtained from K. Goto¹⁵. 35S::PIF4:HA *co-9* was obtained by crossing. The crosses were genotyped for presence of the 35S::PIF4 construct by PCR on genomic DNA using primers 2362 and 2363, resulting in two products of different size representing the complementary DNA (cDNA) transgene and the genomic DNA fragment, respectively. *co-9* was genotyped with primers 3650 and 3652 for insertion, and 3291 and 3292 for the wild-type fragment. For genotyping *phyb-9*, DNA was amplified using oligonucleotides 2137 and 2138 followed by MnlI digestion to distinguish between wild-type and mutant alleles. The global *della* mutant is in the *Ler* background and was described previously⁸. *PIF4::PIF4:ProteinA* and *PIF4::PIF4:GUS* were constructed by amplifying the genomic fragment of *PIF4* including the promoter with oligonucleotides 1534 and 1535. The PCR product was cloned into pENTR/D-TOPO (Invitrogen) and inserted into the binary plasmids PW889 (carboxy-terminal ProteinA) and PW395 (carboxy-terminal GUS), respectively, using Gateway technology (Invitrogen). Transgenic plants were obtained by transforming *pif4-101* by floral dip. For hypocotyl measurements, seeds were surface sterilized, sown on ½ MS media, stratified for 2 days at 4 °C in the dark and germinated for 24 h at 22 °C. The plates were then transferred to short-day conditions (8/16 h photoperiod) at 22 and 27 °C respectively, and grown vertically for 10 days before being imaged and the hypocotyl length measured using the ImageJ software (<http://rsbweb.nih.gov/ij/>). Oligonucleotide sequences are provided in Supplementary Table 1.

Transcript analysis. Samples from plants grown in long days (16/8 h photoperiod) were collected and total RNA was extracted using Trizol Reagent (Invitrogen). RNA (2 µg) was treated with DnaseI (Roche) and used for cDNA synthesis (First strand cDNA synthesis kit, Fermentas). cDNA was diluted 1:8 and used for qPCR with a Roche Lightcycler 480 and the corresponding Sybr Green master mix. To detect *FT* transcript levels, oligonucleotides 3180 and 3181 were used; for *CO*, oligonucleotides 2951 and 2952. *PIF4* transcript levels were analysed using oligonucleotides 3952 and 3953. Oligonucleotides 3247 and 3408 amplifying *TUB6* (At5g12250) were used for normalization.

Immunoblot analysis. To analyse the possible effect of temperature on PIF4 protein stability, plants overexpressing PIF4:HA (35S::PIF4:HA) were used. Seven-day-old 35S::PIF4:HA seedlings grown in short days at 17 °C were transferred to 12, 17, 22 and 27 °C in short days for 2 days. Samples were collected at end of night and

thereafter 30 min, 1 h and 4 h under illumination. Protein samples were separated by SDS–polyacrylamide gel electrophoresis and transferred on to nitrocellulose membrane. PIF4:HA was detected using HRP-conjugated HA antibody (Miltenyi Biotec) and visualized by chemiluminescent detection using Immobilon Chemiluminescent HRP substrate (Millipore).

GUS histochemical assay. For GUS-staining, plants were grown on ½ MS plates in long days (16/8 h photoperiods) for 10 days and kept in the dark for 24 h before collecting. Plants were stained in buffer containing 100 mM phosphate buffer, pH 7, 10 mM EDTA, 0.1% Triton-X100, 0.5 mM K-ferrocyanide and 1 mM X-Gluc at 37 °C for 24 h before de-staining in ethanol.

ChIP. ChIP was performed as described²⁰ with minor modifications. 35S::PIF4:HA seedlings were grown on ½ MS plates for 10 days and kept in the dark for 24 h at respective temperatures before collecting. Plant tissue (1.5 g) and 4 µg of antibody (HA-tag antibody ab9110 from Abcam) were used for ChIP. To analyse the dynamics of *PIF4::PIF4:ProteinA*, plants were grown in respective temperatures under short-day conditions for 4 weeks. Aerial parts of the plants were collected and cross-linked before being used for chromatin preparations. ChIP was done using magnetic beads (Dynabeads M-270 Epoxy, Invitrogen) coated with rabbit IgG (Sigma, I5006) as described (<http://www.ncdir.org/protocols/Rout/Conjugation%20of%20Dynabeads.pdf>). To analyse H2A.Z dynamics at the *FT* locus in response to temperature, we used 3-week-old seedlings of *HTA11::HTA11:GFP* grown at 17 and 27 °C. ChIP was done using anti-GFP antibody (Abcam, ab290). To analyse PIF4 binding in Col-0 and *arp6-1* backgrounds, respective genotypes with *PIF4::PIF4:3XFLAG* were grown on soil at 22 °C under short photoperiods for 3 weeks before samples were fixed by formaldehyde cross-linking. ChIP was performed using anti-Flag M2 affinity gel (Sigma A2220). Immuno-complexes were eluted using 3× Flag peptide (Sigma F4799) according to the manufacturer's instructions. Immunoprecipitated DNA was eluted after reverse cross-linking by boiling at 95 °C for 1 min in the presence of 10% Chelex (BioRad laboratories) followed by treatment with Proteinase K. Oligonucleotides 3255 and 3256 for FT-15 region, 3613 and 3614 for FT-c1, 3607 and 3608 for FT-c and 3261 and 3262 for FT-f were used for detecting PIF4 binding to the *FT* locus. As a positive control for PIF4 binding, At5g45280 was analysed using oligonucleotides 2857 and 2958. *HSP70* was used as a negative control using oligonucleotides 1862 and 1865. To analyse PIF4 binding at At4g39950, oligonucleotides 4240 and 4241 was used for region 1, and oligonucleotides 4246 and 4247 were used for region 2. Oligonucleotides 1860 and 1861 were used for *HSP70* as a negative control. Oligonucleotide sequences are provided in Supplementary Table 1.

Role of corin in trophoblast invasion and uterine spiral artery remodelling in pregnancy

Yujie Cui^{1*†}, Wei Wang^{1*†}, Ningzheng Dong^{2,3*}, Jinglei Lou^{1*}, Dinesh Kumar Srinivasan^{1†}, Weiwei Cheng⁴, Xiaoyi Huang⁴, Meng Liu², Chaodong Fang², Jianhao Peng¹, Shenghan Chen¹, Shannon Wu¹, Zhenzhen Liu², Liang Dong², Yiqing Zhou² & Qingyu Wu^{1,2}

In pregnancy, trophoblast invasion and uterine spiral artery remodelling are important for lowering maternal vascular resistance and increasing uteroplacental blood flow. Impaired spiral artery remodelling has been implicated in pre-eclampsia, a major complication of pregnancy, for a long time but the underlying mechanisms remain unclear^{1,2}. Corin (also known as atrial natriuretic peptide-converting enzyme) is a cardiac protease that activates atrial natriuretic peptide (ANP), a cardiac hormone that is important in regulating blood pressure³. Unexpectedly, corin expression was detected in the pregnant uterus⁴. Here we identify a new function of corin and ANP in promoting trophoblast invasion and spiral artery remodelling. We show that pregnant corin- or ANP-deficient mice developed high blood pressure and proteinuria, characteristics of pre-eclampsia. In these mice, trophoblast invasion and uterine spiral artery remodelling were markedly impaired. Consistent with this, the ANP potently stimulated human trophoblasts in invading Matrigels. In patients with pre-eclampsia, uterine *Corin* messenger RNA and protein levels were significantly lower than that in normal pregnancies. Moreover, we have identified *Corin* gene mutations in pre-eclamptic patients, which decreased corin activity in processing pro-ANP. These results indicate that corin and ANP are essential for physiological changes at the maternal–fetal interface, suggesting that defects in corin and ANP function may contribute to pre-eclampsia.

Pregnancy poses a serious challenge for maintaining normal blood pressure. Pregnancy-induced hypertension, a major cause of maternal and fetal deaths, occurs in approximately 10% of pregnancies^{5,6}. During pregnancy, the uterus undergoes profound morphological changes, including trophoblast invasion and spiral artery remodelling. In pre-eclampsia, impaired spiral artery remodelling is common, but the underlying mechanisms are unclear^{1,2,7–9}. Studies indicate that vascular growth factor receptors, angiotensin and oestradiol are involved in the disease^{10–14}.

Corin is a cardiac protease that activates ANP, which is a cardiac hormone that regulates blood pressure and sodium homeostasis¹⁵. In mice, lack of CORIN prevents ANP generation and causes hypertension¹⁶. In humans, CORIN variants are associated with hypertension¹⁷. Interestingly, *Corin* expression was detected in the pregnant mouse⁴ (Fig. 1A) and human uterus (Supplementary Fig. 1). As a transmembrane protein, CORIN is expected to act at the expression sites, suggesting a possible function in the pregnant uterus.

To understand the role of CORIN in pregnancy, we created a mouse model in which a *Corin* transgene was expressed under a cardiac promoter (Fig. 1B). The transgenic and *Corin* knockout mice were crossed to generate mice expressing *Corin* only in the heart ('knockout/transgenic mice'; Fig. 1C, D). In knockout/transgenic mice, transgenic *Corin* expression restored pro-ANP processing in the heart (Supplementary Fig. 2) and normalized blood pressure (Fig. 1E),

indicating that cardiac CORIN was sufficient to maintain normal blood pressure in non-pregnant mice.

In pregnant *Corin* knockout mice, blood pressure increased at approximately 17 days post coitus and rose further before returning to the non-pregnant blood pressure level after delivery (Fig. 1F), which resembled late gestational hypertension in pre-eclamptic women. In *Corin* knockout/transgenic mice, which were normotensive, blood pressure increased similarly during pregnancy (Fig. 1G), indicating that cardiac *Corin* expression did not prevent pregnancy-induced hypertension. The data also show that in these mice, hypertension in pregnancy was not due to pre-existing high blood pressure. As well as in the uterus, *Corin* mRNA was detected in the umbilical cord and placenta (Supplementary Fig. 3). To distinguish the role of maternal *Corin* from that of placental or other fetal organs, *Corin* knockout females were mated with either wild-type or knockout males. The resulting fetuses carried one or no copy of the functional *Corin* gene. Normally, enzymes that are encoded by one gene copy are able to function. As shown in Fig. 1H, pregnant *Corin* knockout females that were mated with either wild-type or knockout males had similarly increased blood pressure, indicating that lack of maternal, but not fetal, *Corin* caused hypertension in pregnancy.

Proteinuria is a hallmark of pre-eclampsia. Wild-type, *Corin* knockout and knockout/transgenic mice had similar urinary protein levels before pregnancy and at mid gestation. However, the levels increased in *Corin* knockout and knockout/transgenic mice at late gestation (Fig. 1I), consistent with reported proteinuria in mouse models of pre-eclampsia¹⁸. Ischaemic glomeruli, indicated by fewer red blood cells, were found in pregnant *Corin* knockout and knockout/transgenic mice (Fig. 1J, a–f) but not in non-pregnant mice (Supplementary Fig. 4). Periodic acid–Schiff staining revealed increased extracellular matrixes and collapsed glomerular capillaries in pregnant *Corin* knockout and knockout/transgenic mice (Fig. 1J, g–i). Electron microscopy showed narrow glomerular capillary lumens and thick basement membranes (Fig. 1K), suggesting endotheliosis and increased extracellular matrixes. Additional pathological features such as necrotic cells and calcium deposits in the placental labyrinth also existed in these mice (Supplementary Fig. 5), indicating insufficient uteroplacental perfusion. Consistent with this, *Corin* knockout and knockout/transgenic mice had smaller litters (7.1 ± 2.3 ($n = 28$) and 6.8 ± 2.7 ($n = 28$) pups per litter, respectively, versus wild-type mice, which had 9.1 ± 1.2 ($n = 21$) pups per litter; $P < 0.001$ in both cases).

We examined embryos at embryonic day 12.5 (E12.5), an early time point before blood pressure increase in *Corin* knockout and knockout/transgenic mice, and E18.5 (two days before delivery). Wild-type E12.5 embryos showed obvious trophoblast invasion, shown by cytokeratin staining (Fig. 2a), and large vessels mostly in the deep decidua, shown by smooth-muscle α -actin (SMA) staining (Fig. 2b), indicating that

¹Molecular Cardiology, Nephrology and Hypertension, Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, Ohio 44195, USA. ²Cyrus Tang Hematology Center, Jiangsu Institute of Hematology, the First Affiliated Hospital, Soochow University, 199 Ren Ai Road, Suzhou 215123, China. ³Key Lab of Thrombosis and Hemostasis, Jiangsu Institute of Hematology, the First Affiliated Hospital, Soochow University, 188 Shi Zhi Street, Suzhou 215006, China. ⁴The International Peace Maternity and Child Health Hospital, Shanghai Jiaotong University School of Medicine, 910 Hengshan Road, Shanghai 200030, China. [†]Present address: School of Laboratory Science, Tianjin Medical University, Tianjin 300203, China (Y.C.); Department of Cardiology, Peking Union Medical College, Beijing 100730, China (W.W.); Lee Kong Chian School of Medicine, Singapore 637553 (D.K.S.).

*These authors contributed equally to this work.

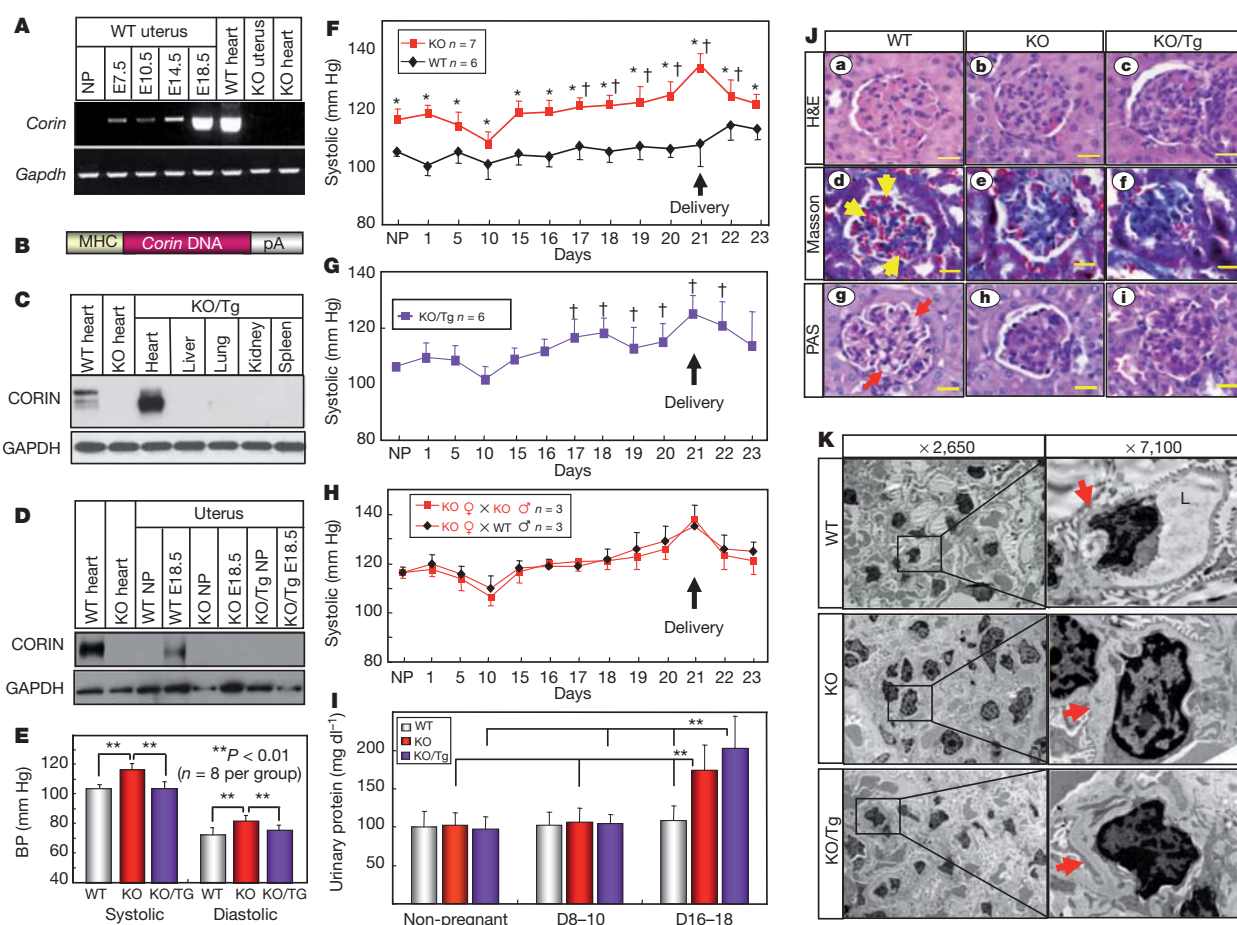


Figure 1 | Hypertension, proteinuria and renal pathology in pregnant *Corin* knockout and knockout/transgenic mice. **A**, *Corin* mRNA expression in mouse uteri. **B**, *Corin* transgenic (Tg) construct. **C**, **D**, Western blot analysis of CORIN protein in wild-type (WT), *Corin* knockout (KO) and knockout/transgenic mice. **E**, Blood pressure (BP, mean \pm s.d.) in non-pregnant females. **F**, **G**, Blood pressure increased in *Corin* knockout (**F**) and knockout/transgenic (**G**) mice in pregnancy. Data are mean \pm s.d. $^{*}P < 0.05$ versus WT of the same time point. $^{\dagger}P < 0.05$ versus non-pregnant level of the same genotype. **H**, Similar blood-pressure changes in *Corin* knockout females mated with knockout or WT males. **I**, Late gestational proteinuria in *Corin*

knockout and knockout/transgenic mice. Data are mean \pm s.d. $^{**}P < 0.01$, $n = 7$ or 8 per group. **J**, **a–i**, Renal ischaemia in pregnant *Corin* knockout and knockout/transgenic mice. E18.5 sections are stained with haematoxylin and eosin (H&E), Masson trichrome or periodic acid–Schiff (PAS). Scale bar, 20 μ m. Red blood cells (yellow arrows) and open capillaries (red arrows) in WT glomeruli are shown. **K**, Narrow glomerular capillary lumen (L) and thick basement membranes (red arrows) in *Corin* knockout and knockout/transgenic mice at E18.5 shown by electron microscopy. GAPDH, glyceraldehyde 3-phosphate dehydrogenase; NP, non pregnant; pA, poly A.

smooth muscles in the superficial decidua were replaced by invading trophoblasts. In contrast, trophoblast invasion in *Corin* knockout and knockout/transgenic embryos was markedly reduced (Fig. 2a) and smaller arteries were found in both superficial and deep decidua (Fig. 2b). In E18.5 wild-type embryos, more abundant trophoblasts were found in the decidua and myometrium compared with those in *Corin* knockout and knockout/transgenic mice (Fig. 2c, d). By haematoxylin and eosin staining, larger and more abundant decidual spiral arteries were observed in wild-type than in *Corin* knockout or knockout/transgenic mice (Fig. 2e). Figure 2f–h shows strong cytokeratin (trophoblasts) staining but weak von Willebrand factor (endothelial) and SMA (smooth muscle) staining in wild-type decidual and myometrial arteries. These data indicate that trophoblast invasion and spiral artery remodelling were impaired in *Corin* knockout and knockout/transgenic mice, and that this defect occurred before blood pressure increased in these mice.

CORIN activates ANP in the heart¹⁵ but it was unknown whether the CORIN function in pregnancy was also mediated by ANP. Pro-ANP is expressed in the non-pregnant and pregnant uterus (Supplementary Fig. 6). If CORIN acts on pro-ANP to promote trophoblast invasion and spiral artery remodelling, thereby preventing hypertension in pregnancy, ANP (also known as *Nppa*) and *Corin* knockout mice should have similar phenotypes. ANP knockout mice are hypertensive (Fig. 3a)

but their blood pressure was not monitored during pregnancy¹⁹. We found similarly increased blood pressure in pregnant ANP knockout mice (Fig. 3b). The mice also had late gestational proteinuria (Fig. 3c) and smaller litters (4.4 ± 1.7 ($n = 25$) versus wild-type, 9.1 ± 1.2 ($n = 21$) pups per litter, $P < 0.001$). By immunostaining, impaired trophoblast invasion and smaller spiral arteries were observed in E12.5 embryos (Fig. 3d, e). In E18.5 embryos, ANP knockout mice had far fewer trophoblasts (Fig. 3f, g) and smaller arteries (Fig. 3h) in the decidua and myometrium than those in wild-type mice. Consistent with this, weak cytokeratin-staining but strong von Willebrand factor-staining were found in arteries in ANP knockout mice (Fig. 3i). Thus, ANP and *Corin* knockout mice had very similar phenotypes, indicating that the role of CORIN in pregnancy is probably mediated by ANP.

In the heart, CORIN produces ANP, which then regulates blood pressure by promoting natriuresis and vasodilation³. Here we found that lack of CORIN and ANP impaired trophoblast invasion and spiral artery remodelling, which was not rescued by cardiac *Corin* expression in *Corin* knockout/transgenic mice. ANP is known to relax vascular smooth muscles. Recently, ANP and its downstream cyclic GMP-dependent protein kinase were shown to be important in angiogenic processes by promoting endothelial regeneration^{20,21}. Thus, ANP may function locally to remodel uterine arteries. Our results also indicate that ANP may directly promote trophoblast invasion (Fig. 4a), and we

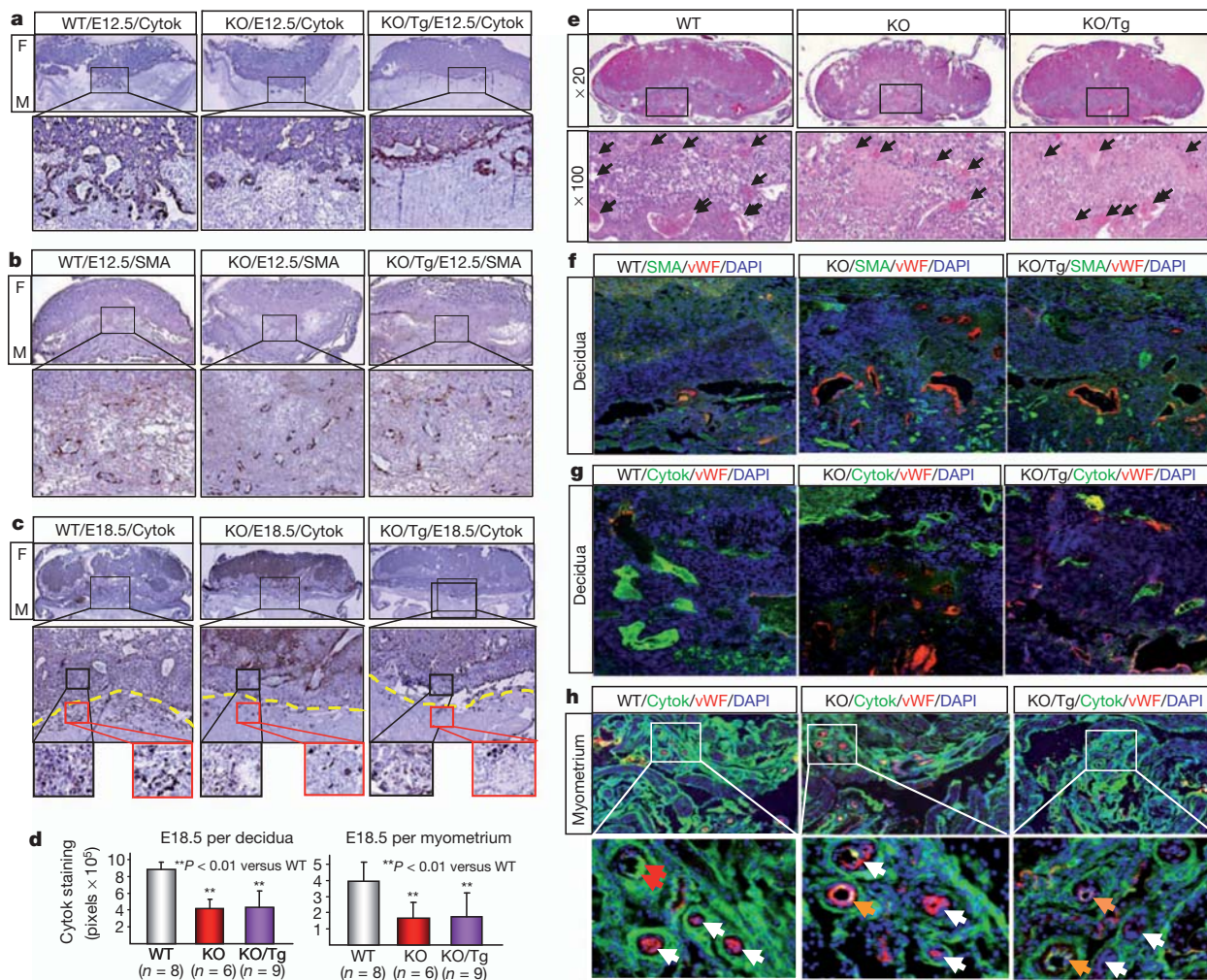


Figure 2 | Impaired trophoblast invasion and spiral artery remodelling in *Corin* knockout and knockout/transgenic mice. **a, b**, E12.5 embryo sections were stained for trophoblasts (**a**) or smooth muscles (**b**). Fetal (F) and maternal (M) sides are indicated. Boxed areas are shown at a higher magnification (×200). **c**, E18.5 embryo sections were stained for trophoblasts. In lower panels (×100), yellow lines show the decidua and myometrium boundary.

therefore tested this idea. We found that ANP markedly stimulated human trophoblasts to invade Matrigels (Fig. 4b) (Supplementary Fig. 7a). In these cells, ANP receptor (also known as atrial natriuretic peptide receptor 1) mRNA expression was confirmed (Supplementary Fig. 7b) and ANP-stimulated intracellular cGMP production was detected (Fig. 4c) (Supplementary Fig. 7c).

Our findings emphasize the importance of local ANP production by CORIN, which acts on trophoblasts and vascular cells in the uterus. Because heart-derived ANP circulates inside the vessel, our model may explain why cardiac CORIN failed to promote trophoblast invasion and uterine artery remodelling, as shown in *Corin* knockout/transgenic mice. To verify this hypothesis, we quantified *Corin* mRNA and protein in human uteruses by polymerase chain reaction with reverse transcription (RT-PCR) and enzyme-linked immunosorbent assay (ELISA). The levels were low in non-pregnant women but increased in pregnant women (Fig. 4d, e). In pre-eclamptic women, the levels were significantly lower than in normal pregnancies. Similar results were found by immunostaining (Fig. 4d and Supplementary Fig. 8). Consistent with this, pro-ANP levels in uterine tissues were significantly higher in pre-eclamptic women than in normal pregnant women (Fig. 4f), indicating that reduced uterine *Corin* expression impaired pro-ANP processing in these patients. *Corin* is a membrane-bound protein^{4,15}, and recent studies showed that CORIN can be shed from

d, Quantitative data (mean ± s.d.) of cytokeratin (Cytok) staining. **e**, Fewer and smaller decidua spiral arteries (arrows) in H&E-stained E18.5 *Corin* knockout and knockout/transgenic embryos. **f–h**, Co-staining of SMA, von Willebrand factor (vWF), cytokeratin and nuclei in E18.5 embryos. Red arrows indicate cytokeratin (green) signals, white arrows indicate von Willebrand factor (red) signals and orange arrows indicate mixed (yellow) signals.

cardiomyocytes and that soluble CORIN is found in human plasma^{22,23}. We found that plasma CORIN levels were higher in pre-eclamptic patients than non-pregnant or normal pregnant women (Fig. 4g). Thus, CORIN levels in plasma did not reflect the levels in tissues, indicating that plasma CORIN was probably derived from the heart, where *Corin* expression increased in response to high blood volume and high blood pressure in pregnancy. These results provide further support for a local function of CORIN in the pregnant uterus.

We next sequenced the *CORIN* gene²⁴ in pre-eclamptic patients and identified a mutation that alters Lys to Glu at position 317 in low-density lipoprotein receptor repeat 2 in one woman (Fig. 4h, j) and another mutation altering Ser to Gly at position 472 in the frizzled 2 domain in two women from the same family who had pre-eclampsia (Fig. 4i, j). In functional studies, Lys317Glu and Ser472Gly mutations did not affect CORIN expression in HEK293 cells but markedly reduced CORIN activity in processing pro-ANP (Fig. 4k–n). The data were consistent with previous findings that Low-density lipoprotein receptor repeats and frizzled domains are critical for CORIN activity²⁵, suggesting that the mutations may impair CORIN function in the patients, thereby contributing to pre-eclampsia. Interestingly, CORIN variants in the frizzled 2 domain that impaired CORIN function have been reported in African American people^{17,26}, a high-risk population for pre-eclampsia.

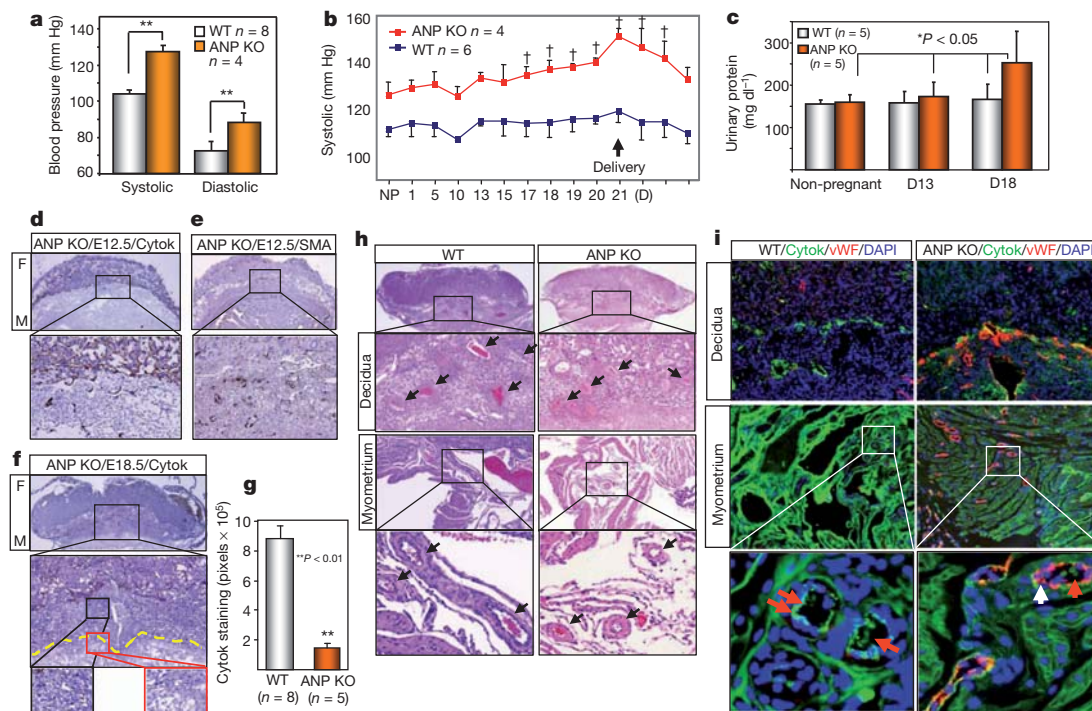


Figure 3 | Hypertension, proteinuria and uteroplacental pathology in pregnant ANP knockout mice. **a**, Blood pressure (mean \pm s.d.) in non-pregnant females, $**P < 0.01$. **b**, Elevated blood pressure (mean \pm s.d.) in pregnant ANP knockout mice. $\dagger P < 0.05$ versus non-pregnant level. **c**, Gestational proteinuria in ANP knockout mice. Data are mean \pm s.d. **d**, **e**, Impaired trophoblast invasion and smooth muscle remodelling in E12.5 embryos stained for cytokeratin (**d**) or SMA (**e**). Boxed areas in top panels are

shown at a higher magnification ($\times 200$). **f**, Impaired trophoblast invasion in E18.5 embryos stained for cytokeratin. **g**, Quantitative data (mean \pm s.d.) of cytokeratin staining in E18.5 ANP knockout embryos. **h**, Impaired decidual and myometrial artery remodelling (arrows) in H&E-stained E18.5 ANP knockout embryos. **i**, Co-staining of cytokeratin, von Willebrand factor and nuclei in E18.5 ANP knockout embryos. Red arrows indicate cytokeratin (green) signals and white arrows indicate von Willebrand factor (red) signals.

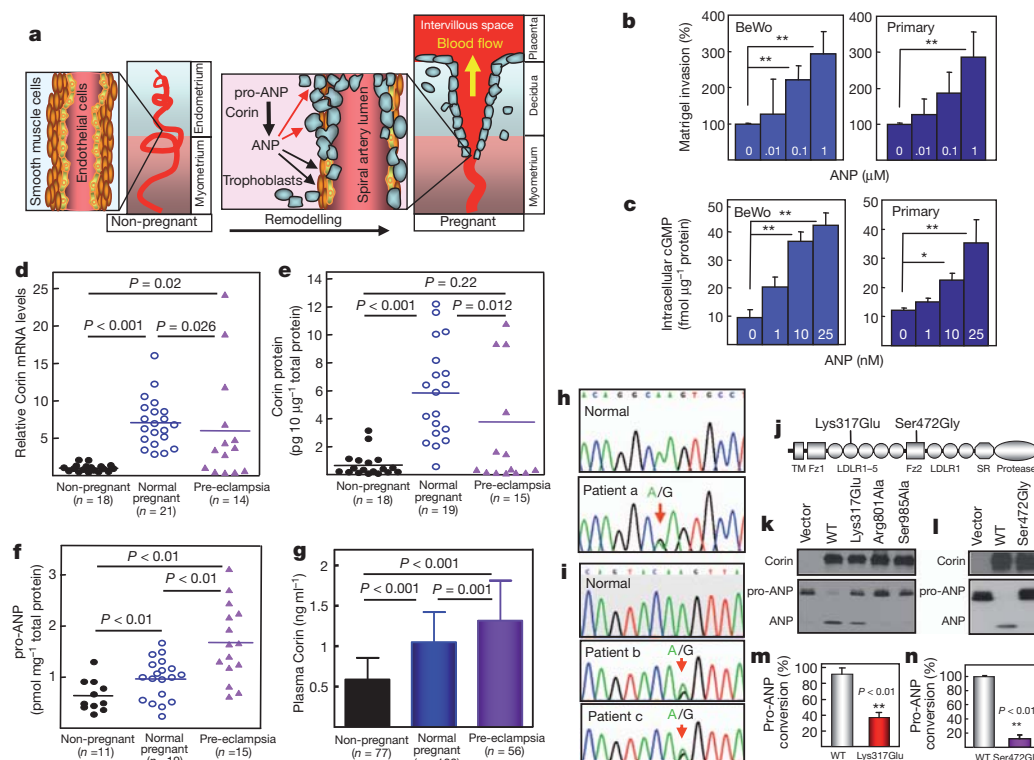


Figure 4 | ANP-stimulated human trophoblast invasion, and impaired uterine Corin expression and Corin mutations in pre-eclamptic patients. **a**, A model showing that CORIN-produced ANP in the pregnant uterus promotes trophoblast invasion (red arrows) and vascular-wall remodelling (black arrows). **b**, **c**, ANP-stimulated human BeWo trophoblasts and primary trophoblasts in Matrigel invasion (**b**) and intracellular cGMP production (**c**). Data are mean \pm s.d. $*P < 0.05$; $**P < 0.01$ versus control. **d–f**, Corin mRNA (**d**) and protein (**e**), and pro-ANP levels (**f**) in human uterus samples. Horizontal lines indicate mean values. **g**, Plasma-soluble CORIN levels (mean \pm s.d.) in pre-eclamptic patients and controls. **h–j**, CORIN gene mutations causing Lys317Glu (**h**) and Ser472Gly (**i**) changes in Corin (**j**). **k**, **l**, Expression of Lys317Glu and Ser472Gly mutants in HEK293 cells (top panels). Vector, WT CORIN and inactive CORIN Arg801Ala and Ser985Ala mutants were controls. Lys317Glu and Ser472Gly mutations reduced pro-ANP processing activity (bottom panels). **m**, **n**, Quantitative data (mean \pm s.d.) from three experiments or more. Fz, frizzled; LDLR, LDL receptor; SR, scavenger receptor; TM, transmembrane.

Previously, high levels of plasma pro-ANP or ANP were detected in pre-eclamptic patients^{27,28}. As shown by our plasma soluble CORIN data, plasma protein levels may not reflect those in tissues. Taken together, our data show a novel local function of CORIN and ANP in promoting trophoblast invasion and spiral artery remodelling to prevent hypertension in pregnancy. The data suggest that impaired *Corin* expression or function in the pregnant uterus may be an important mechanism underlying pre-eclampsia. Studies to better understand impaired uterine *Corin* expression in pre-eclamptic patients may help to develop new strategies to enhance the CORIN-ANP pathway and prevent or treat this life-threatening disease.

METHODS SUMMARY

Corin and ANP knockout mice have been described previously^{16,19}. Transgenic mice with cardiac *Corin* expression were generated using a heart-specific promoter. Blood pressure was measured by radiotelemetry¹⁶. Tissue sections from non-pregnant and pregnant mice were stained with haematoxylin and eosin, Masson's trichrome, periodic acid-Schiff or von Kossa, or immunostained with antibodies against cytokeratin, SMA, von Willebrand factor or CORIN. Renal sections were also examined by electron microscopy. Trans-well invasion assay was carried out with human primary villous trophoblasts (ScienCell) and trophoblastic JEG3, BeWo, JAR cell lines (ATCC) in Matrigel Invasion Chambers (BD Biosciences). ANP-stimulated cGMP production in trophoblasts was assayed in 96-well plates. Intracellular cGMP levels were determined using an enzyme immunoassay kit (Enzo Life Sciences). *Corin* levels in human blood and uterus tissue samples were measured using ELISA²². Pro-ANP levels in human uterus tissues were also measured using ELISA (Alpco Diagnostics). *Corin* gene exons²⁴ from pre-eclamptic patients were PCR-amplified and sequenced directly. *Corin* gene mutations that were identified were studied by expressing mutant CORIN proteins in HEK293 cells and testing their activities in pro-ANP processing assays, as described previously²⁶.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 May 2010; accepted 25 January 2012.

Published online 21 March 2012.

- Pijnenborg, R., Vercruysse, L. & Hanssens, M. The uterine spiral arteries in human pregnancy: facts and controversies. *Placenta* **27**, 939–958 (2006).
- Red-Horse, K. *et al.* Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface. *J. Clin. Invest.* **114**, 744–754 (2004).
- Wu, Q., Xu-Cai, Y. O., Chen, S. & Wang, W. *Corin*: new insights into the natriuretic peptide system. *Kidney Int.* **75**, 142–146 (2009).
- Yan, W., Sheng, N., Seto, M., Morser, J. & Wu, Q. *Corin*, a mosaic transmembrane serine protease encoded by a novel cDNA from human heart. *J. Biol. Chem.* **274**, 14926–14935 (1999).
- Lain, K. Y. & Roberts, J. M. Contemporary concepts of the pathogenesis and management of preeclampsia. *J. Am. Med. Assoc.* **287**, 3183–3186 (2002).
- Sibai, B., Dekker, G. & Kupferminc, M. Pre-eclampsia. *Lancet* **365**, 785–799 (2005).
- Brosens, I. A., Robertson, W. B. & Dixon, H. G. The role of the spiral arteries in the pathogenesis of preeclampsia. *Obstet. Gynecol. Annu.* **1**, 177–191 (1972).
- Kaufmann, P., Black, S. & Huppertz, B. Endovascular trophoblast invasion: implications for the pathogenesis of intrauterine growth retardation and preeclampsia. *Biol. Reprod.* **69**, 1–7 (2003).
- Norwitz, E. R., Schust, D. J. & Fisher, S. J. Implantation and the survival of early pregnancy. *N. Engl. J. Med.* **345**, 1400–1408 (2001).
- Kanasaki, K. *et al.* Deficiency in catechol-O-methyltransferase and 2-methoxyoestradiol is associated with pre-eclampsia. *Nature* **453**, 1117–1121 (2008).
- Levine, R. J. *et al.* Circulating angiogenic factors and the risk of preeclampsia. *N. Engl. J. Med.* **350**, 672–683 (2004).
- Redman, C. W. & Sargent, I. L. Latest advances in understanding preeclampsia. *Science* **308**, 1592–1594 (2005).
- Venkatesha, S. *et al.* Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nature Med.* **12**, 642–649 (2006).
- Zhou, C. C. *et al.* Angiotensin receptor agonistic autoantibodies induce pre-eclampsia in pregnant mice. *Nature Med.* **14**, 855–862 (2008).
- Yan, W., Wu, F., Morser, J. & Wu, Q. *Corin*, a transmembrane cardiac serine protease, acts as a pro-atrial natriuretic peptide-converting enzyme. *Proc. Natl Acad. Sci. USA* **97**, 8525–8529 (2000).
- Chan, J. C. *et al.* Hypertension in mice lacking the proatrial natriuretic peptide convertase *corin*. *Proc. Natl Acad. Sci. USA* **102**, 785–790 (2005).
- Dries, D. L. *et al.* *Corin* gene minor allele defined by 2 missense mutations is common in blacks and associated with high blood pressure and hypertension. *Circulation* **112**, 2403–2410 (2005).
- Davissou, R. L. *et al.* Discovery of a spontaneous genetic mouse model of preeclampsia. *Hypertension* **39**, 337–342 (2002).
- John, S. W. *et al.* Genetic decreases in atrial natriuretic peptide and salt-sensitive hypertension. *Science* **267**, 679–681 (1995).
- Kuhn, M. *et al.* The natriuretic peptide/guanylyl cyclase—a system functions as a stress-responsive regulator of angiogenesis in mice. *J. Clin. Invest.* **119**, 2019–2030 (2009).
- Tokudome, T. *et al.* Impaired recovery of blood flow after hind-limb ischemia in mice lacking guanylyl cyclase-A, a receptor for atrial and brain natriuretic peptides. *Arterioscler. Thromb. Vasc. Biol.* **29**, 1516–1521 (2009).
- Dong, N. *et al.* Plasma soluble *corin* in patients with heart failure. *Circ. Heart Fail.* **3**, 207–211 (2010).
- Jiang, J. *et al.* Ectodomain shedding and autocleavage of the cardiac membrane protease *corin*. *J. Biol. Chem.* **286**, 10066–10072 (2011).
- Pan, J. *et al.* Genomic structures of the human and murine *corin* genes and functional GATA elements in their promoters. *J. Biol. Chem.* **277**, 38390–38398 (2002).
- Knappe, S., Wu, F., Madlansacay, M. R. & Wu, Q. Identification of domain structures in the propeptide of *corin* essential for the processing of proatrial natriuretic peptide. *J. Biol. Chem.* **279**, 34464–34471 (2004).
- Wang, W. *et al.* *Corin* variant associated with hypertension and cardiac hypertrophy exhibits impaired zymogen activation and natriuretic peptide processing activity. *Circ. Res.* **103**, 502–508 (2008).
- Irons, D. W., Baylis, P. H., Butler, T. J. & Davison, J. M. Atrial natriuretic peptide in preeclampsia: metabolic clearance, sodium excretion and renal hemodynamics. *Am. J. Physiol.* **273**, F483–F487 (1997).
- Tihtonen, K. M., Koobi, T., Vuolteenaho, O., Huhtala, H. S. & Uotila, J. T. Natriuretic peptides and hemodynamics in preeclampsia. *Am. J. Obstet. Gynecol.* **196**, 328.e1–328.e7 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Robbins for the α -myosin heavy chain promoter construct and L. Zhang for help with statistical analysis. This work was partly supported by grants from the Ralph Wilson Medical Foundation, the Bakken Heart-Brain Institute and the National Institutes of Health (HL089298, HD064634), and by grants from the National Natural Science Foundation of China (31070716, 81170247 and 31161130356) and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Author Contributions Y.C., W.W., N.D., J.L., D.K.S., M.L., C.F., J.P., S.C., S.W., Z.L. and L.D. designed and performed experiments. N.D., W.C. and X.H. collected patient samples and analysed clinical data. Q.W. conceived the study and designed experiments. Y.Z. and Q.W. wrote the manuscript. All authors analysed and interpreted data, and critically read the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Q.W. (wuq@ccf.org).

METHODS

Knockout and transgenic mice. *Corin* knockout mice were described previously¹⁶. ANP knockout mice (B6.129P2-*Nppa*^{tm1Unc/J})¹⁹ were from the Jackson Laboratory. To make transgenic mice expressing *Corin* in the heart, the full-length mouse *Corin* cDNA was inserted into a construct driven by the mouse α -myosin heavy chain (α -MHC) promoter. Pro-nuclear microinjection and breeding of transgenic mice were carried out at the Case Western Reserve University Transgenic Core. *Corin* knockout and transgenic mice were crossed to generate knockout/transgenic mice. Littermates were used as controls. The animal study was conducted in accordance with the National Institutes of Health guidelines and approved by the Institutional Animal Care and Use Committee at the Cleveland Clinic.

Blood-pressure monitoring. Radiotelemetry was used for real-time blood-pressure monitoring in conscious and unrestrained mice¹⁶. Female mice (8–12 weeks old) were chronically instrumented in the left carotid artery with a PA-C10 device (Data Sciences International) and rested for at least 7 days to recover from the surgery. The mice were mated and checked for vaginal plugs to establish gestation timing. The day on which a plug was observed was defined as E0.5. The mating mice were homozygous except for those in the fetal testing experiment. Telemetry receivers (model RPC-1) were placed under individual cages for data acquisition using the Dataquest A.R.T. 4.0 Gold System (Data Sciences International). Data presented were from continuous recording of at least 6 h per day (10:00 to 16:00).

Urinary protein measurement. Urine samples were collected from non-pregnant mice and pregnant mice at mid (8–10 days post coitus) and late (16–18 days post coitus) gestational stages. Urinary protein levels were measured using a colorimetric assay based on a modified Bradford method (Bio-Rad).

RT-PCR, western blot analysis and ELISA. Total RNAs were isolated from cultured cells or mouse and human tissues using TRIzol reagent (Invitrogen) or an RNeasy kit (Qiagen), and were used to synthesize the first strand cDNAs. RT-PCR was carried out using oligonucleotide primers that were specific for the mouse or human *CORIN*, mouse ANP, or human ANP receptor genes. Quantitative RT-PCR for human *CORIN* mRNA expression in uterus tissues was carried out using the PRISM 7500 System (Applied Biosystems). The β -actin gene was used as an internal control. Quantitative RT-PCR for mouse ANP mRNA in uteruses was carried out using the iCycler system (Bio-Rad). For western blot analysis of *CORIN* protein, membrane fractions from tissue homogenates were isolated by ultracentrifugation, as described previously²⁹. Proteins were analysed by SDS-polyacrylamide agarose gel electrophoresis (SDS-PAGE) and western blot using a polyclonal antibody (Berlex Biosciences). Western blot analysis of pro-ANP in heart samples was carried out using a polyclonal antibody (Santa Cruz). Processing pro-ANP by *CORIN* in transfected cells was analysed by western blot analysis, as described previously³⁰. Pro-ANP in human uterus tissues was measured by an amino-terminal (NT) pro-ANP ELISA kit from Alpco Diagnostics. Human *CORIN* in uterus tissues or plasma was measured by ELISA, as described previously³¹.

Histology and immunohistochemistry. Tissues were fixed with 4% paraformaldehyde and embedded in paraffin. Sections were stained with H&E, Masson's trichrome, PAS or von Kossa. For immunohistochemical or immunofluorescent analysis, antibodies against SMA (Sigma-Aldrich), von Willebrand factor (Sigma-Aldrich) and cytokeratin (Dako) were used to label smooth muscle cells, endothelial cells and trophoblasts, respectively. For human *CORIN*, an antibody from Berlex Biosciences was used. Secondary antibodies were conjugated with horseradish peroxidase or Alexa Fluor 488 (green) or Alexa Fluor 594 (red) (Invitrogen). Tissue sections were mounted with or without DAPI-containing (blue) mounting medium (Dako). For ANP expression in mouse uterus tissues, a polyclonal antibody from Millipore was used. Control sections were treated similarly but without the primary antibodies. Photographs were taken with a light or fluorescent microscope equipped with a digital camera (Olympus). Data are from experiments using five or more mice per study group.

For immunohistochemical analysis of trophoblast invasion in mouse embryos, tissue samples from at least five mice per group, and at least two implant sites per mouse were used. Serial sections (>50 per embryo) of 5 μ m in thickness were prepared. The position of the maternal artery was used as a guide to orient section positions. At least 4–6 sections from the centre of the placenta of each embryo were used for immunohistochemical analysis. Slides that were stained for cytokeratin were examined by two individuals. The sections that showed the deepest trophoblast invasion are presented. These sections were also analysed by ImagePro software to quantify cytokeratin staining. For each section that was analysed, the entire area of the decidua and myometrium was scanned by the software.

Electron microscopy. Kidneys from pregnant mice at E18.5 were fixed in 3% glutaraldehyde, treated with 1% osmium tetroxide and embedded in an Araldite-Epon mixture. Semi-thin sections (0.6 μ m) were prepared and examined with a transmission electron microscope (JEOL JEM-1210) at the Lerner Image Core of the

Cleveland Clinic. Data are from experiments using at least three mice per study group.

Trans-well invasion assay. Human trophoblastic JEG3, BeWo and JAR cells from the American Type Culture Collection were cultured in Minimum Essential Medium (JEG3), RPMI1640 (JAR) and F-12K (BeWo) medium, respectively, with 10% FBS at 37 °C. Primary human villous trophoblasts from ScienCell Research Labs (Carlsbad) were cultured in the Trophoblast Medium (ScienCell) with 10% FBS. Transwell invasion assays were carried out using the BioCoat Growth Factor Reduced Matrigel Invasion Chambers (pore size of 8 μ m) and control inserts (no Matrigel coating) (BD Biosciences) in 24-well plates. Culture medium containing human ANP (Calbiochem) was added to the bottom wells, and cell suspension (5×10^4) was added to the top wells and incubated at 37 °C for 24 h. Non-invading cells were removed from the upper surface of the Matrigel layer by gentle scrubbing. The cells on the lower surface of the membrane were stained using Diff-Quick staining solutions. The membranes were excised and mounted onto glass slides. Invasion indices were determined by counting the number of stained cells on the membrane under a light microscope. The assay was carried out in duplicate in at least three independent experiments.

cGMP assay. ANP-stimulated intracellular cGMP production assay was performed with JEG3, JAR and BeWo cells and primary human trophoblasts using a method described previously³². The cells were grown in 96-well plates. Confluent cells were washed once with serum-free medium. Human ANP was added to serum-free medium and incubated with cells at 37 °C for 30 min. In these experiments, ANP was more potent in stimulating intracellular cGMP production when serum-free medium was used (data not shown). The cells were lysed with 0.1 M HCl. Intracellular cGMP levels in ANP-stimulated cells were determined using an EIA kit (Enzo Life Sciences). Each experimental condition was assayed in duplicate in at least three independent experiments.

Human blood and tissue samples. The study was approved by local ethics committees and participants gave informed consent. Women of normal pregnancy or with pre-eclampsia, and age-matched non-pregnant normal controls who underwent routine medical check-ups were recruited. All participants were ethnic Han Chinese. Hypertension was defined as diastolic pressure >90 mm Hg and/or systolic pressure >140 mm Hg on at least two occasions. Pre-eclampsia was defined as hypertension that appeared after 20 weeks of gestation with proteinuria (>300 mg urinary protein per 24 h). Patients with chronic hypertension, chronic kidney disease, diabetes and heart disease were excluded. Uterus tissues were obtained during caesarean sections in pregnant women or operations for uterine leiomyoma in non-pregnant women. Clinical characteristics of women who provided blood and those who provided uterus tissue samples are summarized in Supplementary Tables 1 and 2, respectively.

***CORIN* gene sequences in patients.** Blood samples from 56 patients with pre-eclampsia were collected into tubes containing EDTA as an anticoagulant. Genomic DNA was extracted from white blood cells using the QIAamp DNA Mini kit (Qiagen) and used in PCR to amplify exon sequences of the *CORIN* gene²⁴. PCR products were used for direct DNA sequencing. Mutations that were identified were verified by independent PCR and DNA sequencing. Additional PCR and DNA sequencing were carried out with DNA samples from more than 100 normal controls to verify that mutations that were identified in patients did not exist in the normal population.

Expression and functional analysis of *Corin* mutants. Plasmids expressing human wild-type *Corin* and two inactive mutants Arg801Ala and Ser985Ala, in which the activation cleavage site and catalytic site residues were mutated, respectively, were described previously³³. Plasmids expressing *Corin* mutants Lys317Glu or Ser472Gly were constructed by PCR-based mutagenesis. Recombinant *CORIN* proteins that were expressed by these plasmids contained a carboxy-terminal V5 tag to be detected by an anti-V5 antibody (Invitrogen)³⁰. Plasmids were transfected into HEK293 cells using Lipofectamine 2000 (Invitrogen). Cells were lysed and proteins were analysed by western blot using an anti-V5 antibody. To analyse the function of *CORIN*, recombinant human pro-ANP in conditioned medium was added to HEK293 cells expressing *Corin* wild-type or mutants and incubated at 37 °C for 2 h. Pro-ANP and ANP in the medium were immunoprecipitated and analysed by western blot. Protein bands on X-ray films were scanned by densitometry. The percentage of pro-ANP to ANP conversion was calculated as described previously³⁰.

Statistical analysis. Results are presented as mean \pm s.d. Differences between two groups were analysed with the Student's *t*-test. Data involving more than two groups were analysed by analysis of variance followed by the Tukey multiple comparison test. Comparisons for *Corin* mRNA and protein and pro-ANP levels in human uterus samples were carried out using the Mann-Whitney-Wilcoxon test. A *P* value of less than 0.05 was considered statistically significant.

29. Chen, S. *et al.* Protease corin expression and activity in failing hearts. *Am. J. Physiol. Heart Circ. Physiol.* **299**, H1687–H1692 (2010).

30. Liao, X., Wang, W., Chen, S. & Wu, Q. Role of glycosylation in corin zymogen activation. *J. Biol. Chem.* **282**, 27728–27735 (2007).
31. Dong, N. *et al.* Effects of anticoagulants on human plasma soluble corin levels measured by ELISA. *Clin. Chim. Acta* **411**, 1998–2003 (2010).
32. Wu, F., Yan, W., Pan, J., Morser, J. & Wu, Q. Processing of pro-atrial natriuretic peptide by corin in cardiac myocytes. *J. Biol. Chem.* **277**, 16900–16905 (2002).
33. Qi, X., Jiang, J., Zhu, M. & Wu, Q. Human corin isoforms with different cytoplasmic tails that alter cell surface targeting. *J. Biol. Chem.* **286**, 20963–20969 (2011).

Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function

Colin R. Lickwar^{1*}, Florian Mueller^{2,3†*}, Sean E. Hanlon^{1†}, James G. McNally² & Jason D. Lieb¹

Dynamic access to genetic information is central to organismal development and environmental response. Consequently, genomic processes must be regulated by mechanisms that alter genome function relatively rapidly^{1–4}. Conventional chromatin immunoprecipitation (ChIP) experiments measure transcription factor occupancy⁵, but give no indication of kinetics and are poor predictors of transcription factor function at a given locus. To measure transcription-factor-binding dynamics across the genome, we performed competition ChIP (refs 6, 7) with a sequence-specific *Saccharomyces cerevisiae* transcription factor, Rap1 (ref. 8). Rap1-binding dynamics and Rap1 occupancy were only weakly correlated ($R^2 = 0.14$), but binding dynamics were more strongly linked to function than occupancy. Long Rap1 residence was coupled to transcriptional activation, whereas fast binding turnover, which we refer to as ‘treadmilling’, was linked to low transcriptional output. Thus, DNA-binding events that seem identical by conventional ChIP may have different underlying modes of interaction that lead to opposing functional outcomes. We propose that transcription factor binding turnover is a major point of regulation in determining the functional consequences of transcription factor binding, and is mediated mainly by control of competition between transcription factors and nucleosomes. Our model predicts a clutch-like mechanism that rapidly engages a treadmilling transcription factor into a stable binding state, or vice versa, to modulate transcription factor function.

The diverse biological functions of Rap1 (ref. 9) make it an excellent model for testing the hypothesis that binding dynamics are important for transcription factor function (Supplementary Fig. 1). We developed a strain with two copies of *RAP1*. One copy of *RAP1* was tagged with a 3× Flag epitope and was constitutively expressed from the endogenous *RAP1* promoter. A second copy of *RAP1* was tagged with a 9× Myc epitope and was controlled by a weakened galactose-inducible promoter, *GALL* (an attenuated version of the *GAL1* promoter) (Fig. 1a). This strain showed no growth defects in inducing (2% galactose) or non-inducing (2% dextrose) conditions (Fig. 1b and Supplementary Fig. 2). To avoid cell-cycle and DNA replication effects, for the duration of the experiment the strain was arrested in G1 with alpha factor⁶. The induced Rap1 protein isoform could be detected as early as 30 min after galactose induction (Fig. 1c). The ratio of Rap1 isoforms provided an estimate of the nucleoplasmic pool of Rap1 molecules (Fig. 1d). We then performed Myc and Flag ChIP experiments independently from extract corresponding to each of 10 time points (0, 10, 20, 30, 40, 50, 60, 90, 120 and 150 min after induction). We also performed ChIP to measure total Rap1 occupancy using a Rap1-specific antibody at 0 and 60 min. DNA fragments enriched in the

ChIPs were detected on whole-genome-tiling 12-plex microarrays containing 270,000 probes per subarray, with an average probe interval of 41 bp and an average probe length of 54 bp (Supplementary Fig. 3). The entire time-course experiment was performed in duplicate. (Procedural details can be found in the Methods.)

After induction, Rap1–Myc was incorporated at targets where Rap1 had previously been shown to bind^{8,10} (Fig. 2a, b), indicating that the system was functioning as designed. The increase in Rap1 protein caused by the induction of the competitor did not cause an increase in the overall occupancy at the measured Rap1 sites (Fig. 2c, d and Supplementary Figs 4 and 5). As Rap1–Myc ChIP occupancy increased at sites of Rap1 binding, Rap1–Flag occupancy decreased coordinately (Fig. 2c, d and Supplementary Fig. 4). Thus, Rap1–Myc is competing specifically with Rap1–Flag at each locus, and Rap1–Myc binding is not the result of cooperativity or additional Rap1 binding locations.

To interpret our data, we developed a model to determine turnover rates of Rap1 by modifying a fitting algorithm used previously to

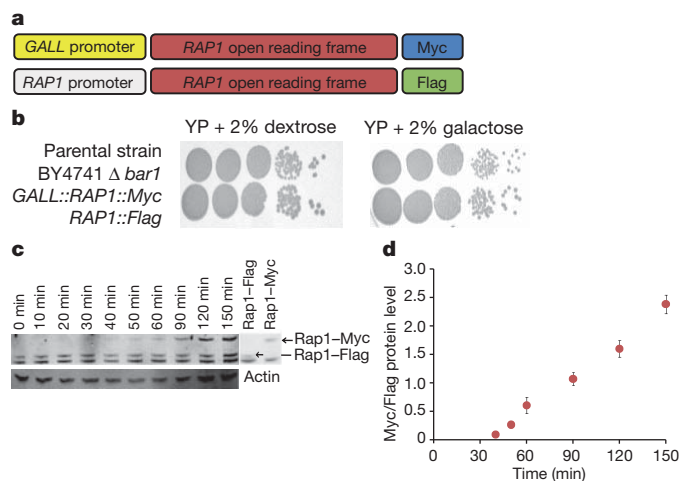


Figure 1 | Development of transcription factor competition ChIP yeast strains. **a**, Schematic of the Rap1 competition-ChIP yeast strain. **b**, Growth comparison of competition yeast strain and wild-type in inducing (2% galactose) and non-inducing (2% dextrose) conditions. **c**, Western blot analysis using an antibody against Rap1 (y-300). Strains containing only a Rap1–Myc or only Rap1–Flag copy are shown to the right to indicate the size of isoform-specific bands. The actin loading control is shown below. **d**, To estimate the dynamics of induction, the ratio of induced Rap1–Myc and constitutive Rap1–Flag protein is plotted. Data are from two technical replicates of two independent time-course replicates. Error bars represent s.e.m.

¹Department of Biology, Carolina Center for the Genome Sciences, Curriculum in Genetics and Molecular Biology, and Lineberger Comprehensive Cancer Center, CB 3280, 408 Fordham Hall, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA. ²LRBGE-National Cancer Institute, The National Institutes of Health, 41 Library Drive, Bethesda, Maryland 20892, USA. ³Institut Pasteur, Groupe Imagerie et Modélisation, Centre National de la Recherche Scientifique, Unité de Recherche Associée 2582, 25-28 rue du Docteur Roux, 75015 Paris, France. [†]Present addresses: Institut de Biologie de l'Ecole Normale Supérieure, Functional Imaging of Transcription, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 8197, 45 rue d'Ulm, 75005 Paris, France (F.M.); Office of Physical Sciences-Oncology, CSSI, NCI, NIH, 31 Center Drive, Room 10A03, Bethesda, Maryland 20892, USA (S.E.H.).

*These authors contributed equally to this work.

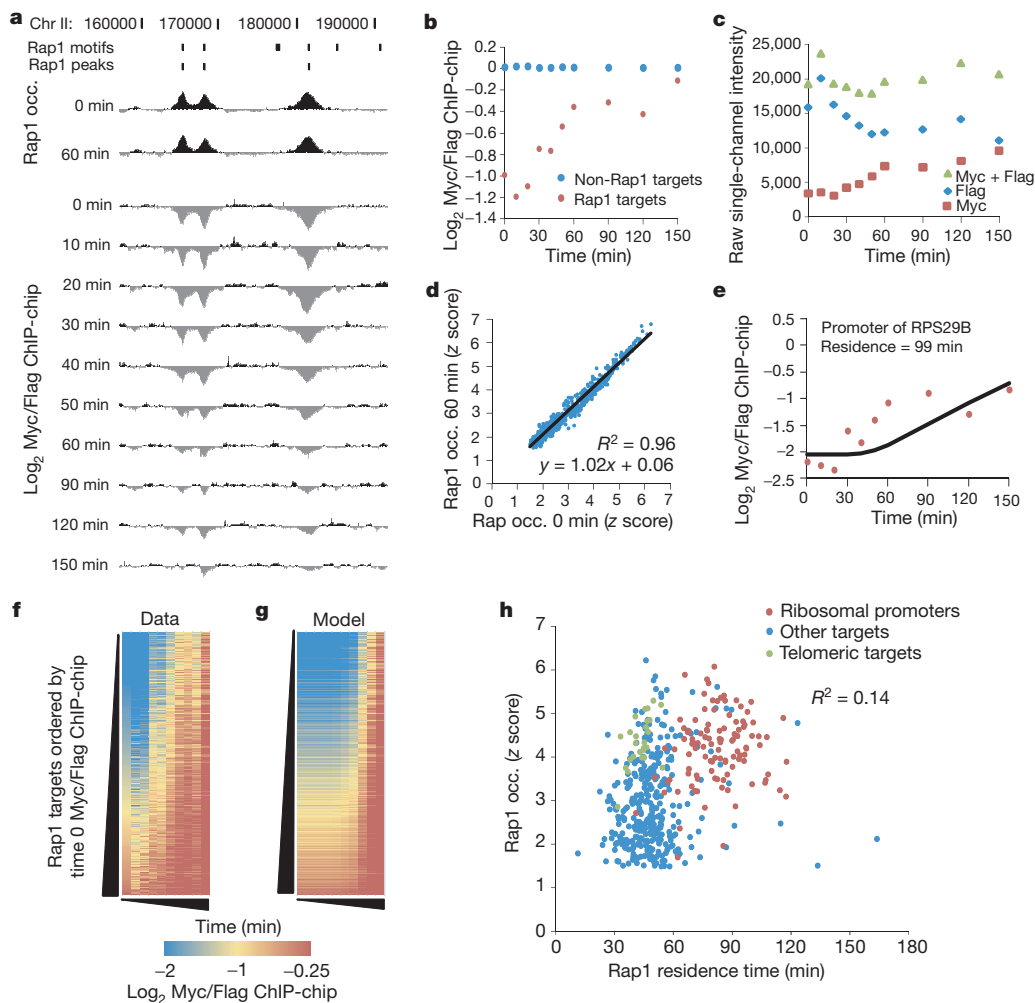


Figure 2 | Rap1-bound sites exhibit distinct replacement dynamics. **a**, A Rap1 turnover experiment over a 30-kb region of chromosome II (Chr II). Rap1 motifs and peaks are shown. **b**, Average \log_2 Myc/Flag ratio values for all Rap1 targets (red) increase relative to non-Rap1 targets (blue). **c**, Rap1–Myc competes with Rap1–Flag for binding. Average single-channel intensity for Rap1–Myc and Rap1–Flag for a single probe (CHR15F000978891) in the promoter of the *TYE7* gene shows that the increase in Rap1–Myc is coincident with the loss of Rap1–Flag. **d**, Total Rap1 occupancy does not change during the time course. Average total Rap1 occupancy (occ.) (\log_2 ratio of Rap1 immunoprecipitation

(y-300) to input ratio z score) at Rap1 targets at time 0 is plotted against that at 60 min. **e**, Average \log_2 Myc/Flag ratios for the promoter of ribosomal protein gene *RPL29B* (red points). The model fit for the residence time parameter that best fits these data is shown (black line). **f**, Colorimetric representation of \log_2 Myc/Flag ratios for all 465 Rap1 targets, sorted by the initial (normalized) \log_2 Myc/Flag ratio. **g**, For each site in **f**, the \log_2 Myc/Flag ratio predicted by our residence time model, based on the calculated residence time, is shown. **h**, Rap1 occupancy (time 0 z score) versus Rap1 residence for 465 Rap1 targets ($R^2 = 0.14, 0.37$; Spearman rank correlation).

measure histone H3 turnover⁶. Under our experimental conditions, the extracted turnover rate for a transcription factor at a binding site is equivalent to its dissociation rate, which allows us to measure the residence time (Supplementary Figs 6–8 and Supplementary Text). Our experimental system can quantify binding events that have an apparent duration of 500 s or longer (Supplementary Figs 6 and 7, and Supplementary Text). Using our ChIP data and model we measured residence time of Rap1 at 439 peaks of Rap1 enrichment genome-wide, and at the 26 uniquely mappable telomeres (Fig. 2e–h, Supplementary Fig. 9 and Supplementary Text). Rap1 occupancy correlated only modestly with Rap1 residence ($R^2 = 0.14$; Spearman rank correlation = 0.37) (Fig. 2h), and distinct dynamics of Rap1–Myc incorporation were observed at different genomic loci (Fig. 2e–h). Thus, residence time and occupancy are distinct measurements, and our system was capable of distinguishing Rap1 turnover kinetics at different loci in the same experiment.

We found that efficient transcriptional activation was associated with stable Rap1 binding, whereas lower transcript production was associated with treadmilling, despite similar levels of Rap1 occupancy. Long Rap1 residence times occurred at ribosomal protein gene

promoters, which are very highly transcribed and strongly activated by Rap1 (refs 6, 11, 12) (Fig. 2e, h). In contrast, Rap1 binding to non-ribosomal protein targets and to the infrequently transcribed telomeric and subtelomeric Rap1 sites was characterized by fast turnover (Fig. 2h and Supplementary Figs 10 and 11). Stable Rap1 binding seems to support higher mRNA production through more efficient recruitment of the RNA polymerase II machinery¹² (Fig. 3a). Genes with stable Rap1 binding at their promoters indeed exhibited high levels of RNA polymerase II association⁶ (Fig. 3a), high transcription initiation rates^{11,13} (Fig. 3b) and high mRNA levels (Fig. 3c). Rap1 occupancy does not correspond as strongly (Fig. 3a–c; right panels). TATA-binding protein (TBP) turnover⁷ is also slow at ribosomal protein genes, indicating that slow transcription-factor-binding dynamics may be a hallmark of efficient transcription initiation^{11,13}.

We next examined possible mechanisms for the locus-specific differences in Rap1 residence time. Nucleosomes are a major regulator of genome accessibility¹⁴, so we examined the relationship between histone modification and Rap1-binding dynamics^{15,16}. Sites of long Rap1 residence were strongly correlated with sites of enrichment for the histone acetyltransferases Gcn5 and particularly Esa1 (ref. 17) (Fig. 3d, e).

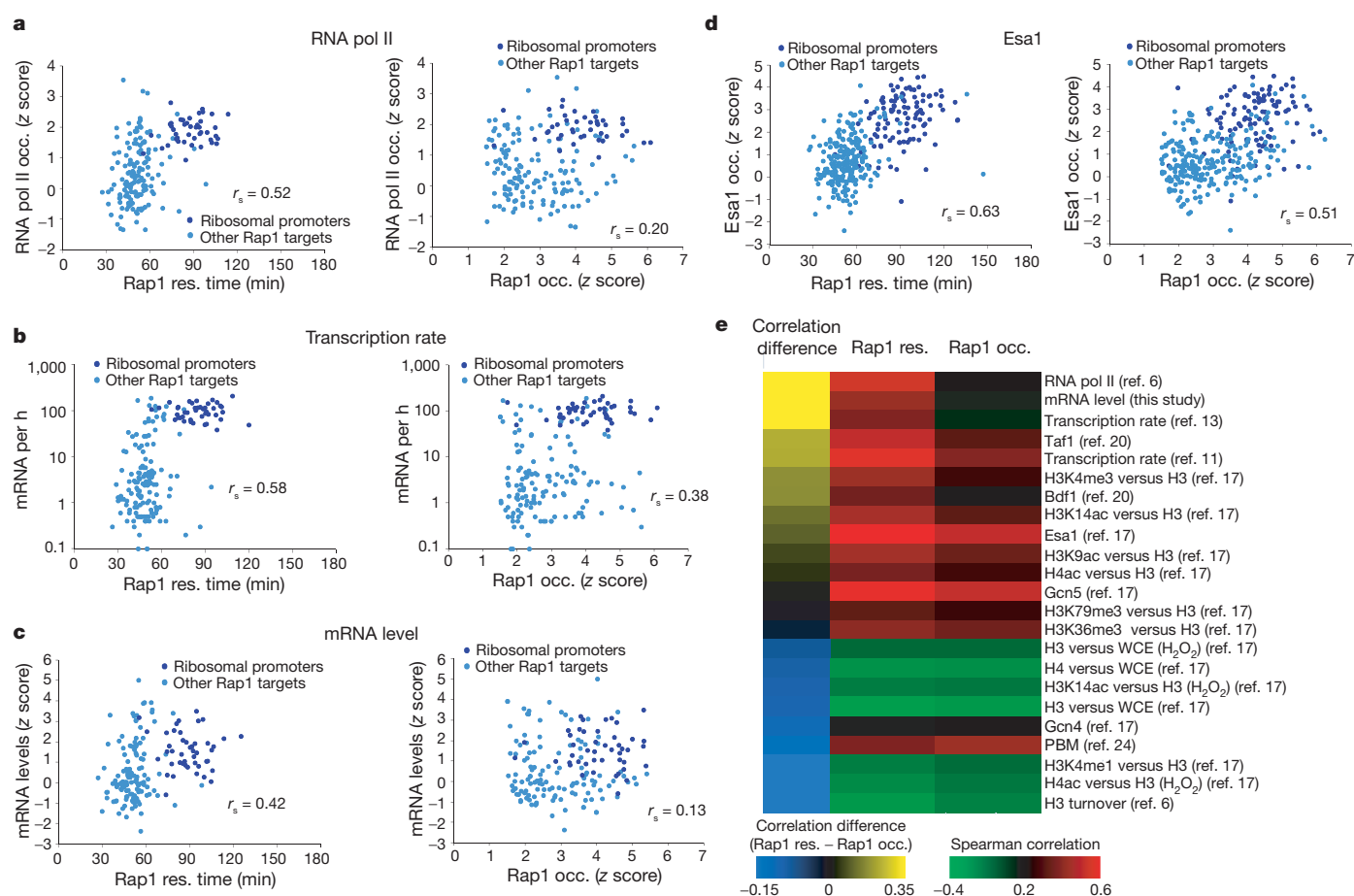


Figure 3 | RNA polymerase II recruitment, mRNA production, and histone acetyltransferase recruitment is associated with long Rap1 residence.

a–d, Left, Rap1 residence (res.) time is plotted on the *x* axis. Right, Rap1 occupancy is plotted on the *x* axis. In both panels, the following are plotted on the *y* axis: RNA polymerase II (RNA Pol II) occupancy (**a**)⁶; the number of mRNA transcripts per h (**b**) (ref. 11); mRNA levels at time 0 (**c**); and histone acetyltransferase Esa1 occupancy z scores (**d**)¹⁷. r_s is the Spearman correlation

Nucleosome instability reinforced by Gcn5 and Esa1 (members of SAGA and NuA4, respectively) may stabilize Rap1 binding by reducing competition with nucleosomes^{18,19}. Other indicators of active promoters—including H3K4me3, occupancy by the bromodomain protein Bdf1 (similar to mammalian TAF1)²⁰, and acetylation of H3K9, H4 and H3K14—were also more strongly associated with Rap1 residence time than with Rap1 occupancy (Fig. 3e).

In general, sites that are bound by Rap1 are strongly depleted of nucleosomes²¹. However, the binding dynamics data allowed us to appreciate a more complex relationship. We grouped Rap1-bound loci into four categories based on their measured Rap1 residence time: longest, long, short, and shortest. We then aligned the Rap1 motifs in each category and plotted nucleosome occupancy relative to the motif position, reasoning that nucleosomes in direct proximity to the DNA motif bound by Rap1 would have a strong influence on Rap1 residence times²¹. As expected, strong nucleosome depletion was centred on the Rap1 motif (Fig. 4a). However, as Rap1-binding turnover increased, nucleosome depletion was correspondingly less pronounced. Thus, not all highly occupied Rap1 sites are equally depleted of nucleosomes *in vivo*. Instead, a subset of loci at which Rap1 occupancy is high but binding turnover is also high (treadmilling) are associated with higher nucleosome occupancy (Supplementary Fig. 11b). No consistent relationship is apparent when Rap1 targets are grouped by occupancy, as measured by traditional ChIP (Fig. 4a).

We next examined nucleosome occupancy on naked DNA in the absence of Rap1 or any protein cofactors²¹. Notably, DNA-encoded

value. **e**, Colorimetric representation of the Spearman rank correlation between various genomic data sets and Rap1 occupancy (right) and Rap1 residence (centre), ordered by the magnitude of the difference between the residence and occupancy correlations for each comparison (left). WCE, whole-cell extract; PBM, protein-binding microarray. Telomeric targets are excluded from analysis (Supplementary Text).

nucleosome occupancy measured *in vitro* is low only for the class of Rap1 targets with the most stable binding (Fig. 4b and Supplementary Fig. 11b). This pattern was not recapitulated when Rap1 targets were sorted by occupancy (Fig. 4b). This suggests that the nucleosome behaviour surrounding transcription factor motifs is at least partially encoded in DNA¹⁵, and that this DNA-encoded nucleosome occupancy can influence the binding dynamics of transcription factors, and thereby affect functional outputs (Supplementary Fig. 11a–c).

We sought further evidence supporting direct competition between nucleosomes and Rap1. We compared histone H3 turnover⁶ to Rap1 residence times and found that loci with long Rap1 residence times also had relatively slow H3 turnover. Similarly, histone H3 molecules that treadmill are found almost exclusively at sites of Rap1 treadmill (Fig. 4c). Rap1–nucleosome interactions isolated by immunoprecipitating Rap1 after MNase digestion²² were also detected more often at treadmill sites (Fig. 4d). Further evidence for competition is supported by a marked increase in nucleosome occupancy directly over Rap1 motifs after Rap1 depletion²³ at treadmill loci, but not at loci with stable Rap1 binding (Fig. 4e). These relationships provide evidence for direct competition between Rap1 and nucleosomes.

Given that high DNA-encoded nucleosome occupancy is associated with rapid Rap1 turnover (Fig. 4b), it is reasonable to expect that differences in the strength of the DNA motif bound by the transcription factor would also influence turnover. To test this, we examined the relationship between Rap1 turnover and experimentally measured *in vitro* Rap1 affinity at each locus²⁴. For sites with longer Rap1 residence,

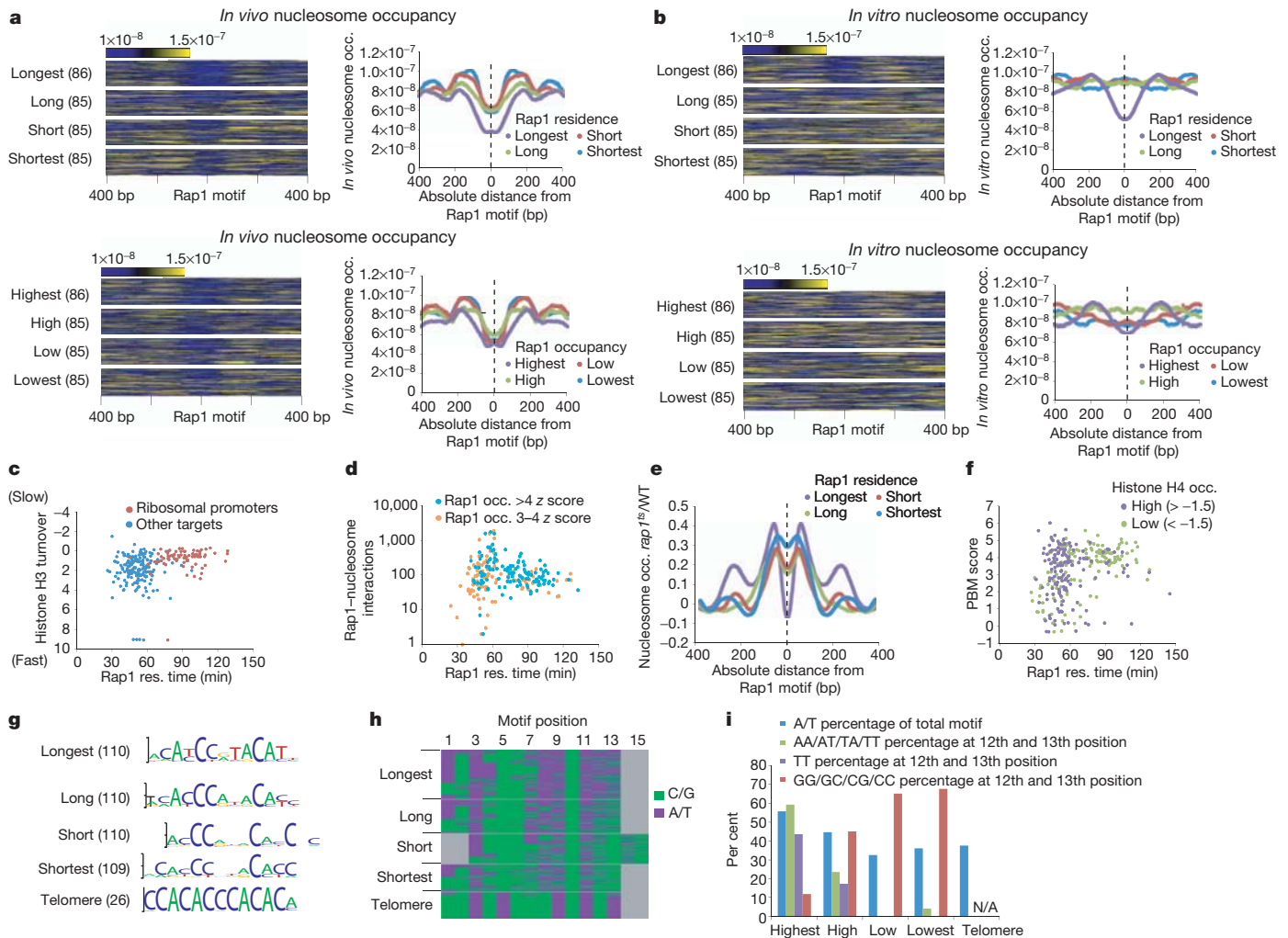


Figure 4 | Evidence for competition between Rap1 and nucleosomes.

a, Colorimetric representation of *in vivo* nucleosome occupancy centred on Rap1-binding motifs. Loci are ordered by Rap1 residence time (top) or Rap1 occupancy (bottom). The total number of Rap1 targets in each group is shown in parentheses. To the right are plots of the average nucleosome occupancy for each group centred on the Rap1 motif. Targets with several Rap1 motifs are represented by one randomly chosen motif. **b**, Same as (a) but for *in vitro* nucleosome occupancy. **c**, Histone H3 turnover versus Rap1 residence for ribosomal protein genes (red) and other targets (blue). **d**, The number of Rap1-nucleosome interactions²² detected within each Rap1 target peak boundary on a log₁₀ scale. **e**, Relative change in nucleosome occupancy following Rap1 depletion²³ centred on Rap1 motifs grouped by residence time. A value of zero

represents no relative change in nucleosome occupancy. **f**, *In vitro* Rap1 affinity for its cognate target, as measured using a PBM²⁴, compared to Rap1 residence. Colours represent histone H4 occupancy z scores (> -1.5 purple (high), < -1.5 green (low)¹⁷). **g**, Top-position weight matrix motifs discovered for Rap1 targets, grouped by residence. The number of targets for each group is in parentheses. **h**, All motifs from the top-position weight matrix for each residence group are coloured by their A or T (A/T) (purple), or G or C (G/C) (green) content at each motif base position. **i**, Percentage of A/T content for the entire motif (blue), AA/AT/TA/TT at the twelfth and thirteenth motif position (green), TT at the twelfth and thirteenth position (purple) and GG/GC/CG/CC at the twelfth and thirteenth position (red) for Rap1 targets, grouped by residence and telomeric regions.

Rap1's affinity for DNA was generally high, whereas Rap1 sites with the fastest turnover had lower experimentally measured Rap1 affinity (Fig. 4f). Despite this relationship, among sites with strong Rap1 motifs, nucleosome occupancy was still the major factor distinguishing sites with long Rap1 residence times from those with higher turnover (Fig. 4f).

Longer *in vivo* Rap1 residence times at sites of high Rap1 affinity are consistent with control of the Rap1-nucleosome competition being encoded directly in the DNA sequence to a substantial degree^{15,16}. We reasoned that this would be reflected in the sequence of the DNA motifs bound by Rap1. Indeed, we found differences in the composition of the Rap1 motifs for each of the turnover categories, with the longest residence Rap1 sites preferentially containing A or T at positions 4, 8, 12 and 13 (Fig. 4g-i). These associations were not as strong when Rap1 targets were ordered by occupancy (Supplementary Fig. 12). Sites at which residence was shortest tended to contain a degenerate Rap1 binding motif (Fig. 4g).

For several other transcription factors, microscopy-based measurements at individual loci point to much shorter residence times than those measured for Rap1 (refs 1, 3, 4, 25-27). For example, despite an *in vitro* residence time similar to Rap1 (~90 min^{9,28}), glucocorticoid-receptor binding seems to be exceptionally short-lived at individual loci^{26,27}. Nevertheless, an overall positive relationship between residence time and transcriptional output is observed for both Rap1 and glucocorticoid receptor^{3,29}. The differences in Rap1 and glucocorticoid-binding dynamics, and the disparity between glucocorticoid-receptor residence time *in vitro* and *in vivo*, may reflect different modes of interactions with nucleosomes. The binding affinity of glucocorticoid receptor may be particularly sensitive to nucleosome packaging or may be regulated by the availability of DNA that is transiently accessible from the nucleosome surface¹⁸. This type of accessibility on the nucleosome itself could be regulated, and would not rely on the complete loss of a nucleosome¹⁴. Rap1 itself exhibits such properties, with its binding progressively inhibited as the motif recognized by

Rap1 is moved closer to the nucleosome dyad³⁰. Our data do not exclude a model in which transcription factor binding occurs adjacent to a nucleosome, and competition occurs without complete nucleosome eviction^{14,22,30}.

In this study, we determined Rap1-binding dynamics genome-wide using competition ChIP. Rap1 occupancy was only weakly correlated with Rap1 binding turnover, showing that these are independently measurable properties. Binding turnover correlates more strongly than occupancy with many aspects of genomic function, most predominantly RNA polymerase II recruitment and transcript levels. Stable Rap1 binding is associated with activation, whereas Rap1 treadmilling is associated with higher nucleosome occupancy, nucleosomal treadmilling, and a lack of transcription. Our work provides the basis for a model in which transcription-factor-binding dynamics is a major point of regulation in determining the functional consequences of transcription factor binding. Importantly, this model provides a plausible mechanism for a locus-specific switch between inactive and active transcription factor states, or even for a rapid switch from an activator (stable binding) to a repressor (treadmilling), or vice versa. This could be achieved at any given locus through a 'clutch' that alters the balance of the continual competition between transcription factors and nucleosomes (Supplementary Fig. 1). This clutch could operate through histone modification, histone-variant incorporation, ATP-dependent chromatin remodelling, cofactor binding, or any other site-directed chromatin altering activity.

METHODS SUMMARY

Strain construction. The *RAP1* gene and promoter was cloned into the pRS403 plasmid and integrated by homologous recombination into the *HIS3* locus of the BY4741 *S. cerevisiae* strain. The two copies of *RAP1* were then sequentially tagged using the 9× Myc epitope from pYM20:hphNT1 at the *HIS3* copy of *RAP1* and the 3× Flag tag from p3Flag-KanMX at the endogenous *RAP1* copy. The *HIS3* copy of the *RAP1* promoter was replaced using homologous recombination by amplifying the *GALL:natNT2* promoter from the pYM-N27 plasmid. Integrations were confirmed using PCR and western blot analysis. The *BARI* gene was knocked out by homologous recombination using a *LEU2* gene amplified from pRS405.

Time course. Yeast were grown overnight in YPD (yeast extract 1%, peptone 2% and dextrose 2%) and used to inoculate 800 ml of YPR (yeast extract 1%, peptone 2% and raffinose 2%) to an attenuation at 600 nm ($D_{600\text{ nm}}$) of 0.2 (Genesys 20 Spectrophotometer) in a 4-l Erlenmeyer flask. These cells were grown to an D_{600} of 0.4 and subsequently arrested using 5 μ M alpha factor (400 μ l of 10 mM, GenScript) until 95% of the yeast cells were unbudded (~3 h). Cells were then induced by adding 40% galactose to a final concentration of 2%. At this time, additional alpha factor was added (400 μ l of 10 mM, GenScript). Samples were collected at time points 0, 10, 20, 30, 40, 50, 60, 90, 120 and 150 min after galactose induction. At each time point, 35 ml of culture was taken and added immediately to 37% formaldehyde to a final concentration of 1% for 20 min. Thirteen millilitres were taken for subsequent RNA preparation. Two millilitres were taken for protein preparation by pelleting cells and heating at 95 °C for 5 min in 0.06 M Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% 2-mercaptoethanol and 0.0025% bromophenol blue. All samples were frozen immediately in liquid nitrogen.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 February 2011; accepted 23 February 2012.

- Mueller, F., Wach, P. & McNally, J. G. Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophys. J.* **94**, 3323–3339 (2008).
- Yao, J., Munson, K. M., Webb, W. W. & Lis, J. T. Dynamics of heat shock factor association with native gene loci in living cells. *Nature* **442**, 1050–1053 (2006).
- Stavreva, D. A., Muller, W. G., Hager, G. L., Smith, C. L. & McNally, J. G. Rapid glucocorticoid receptor exchange at a promoter is coupled to transcription and regulated by chaperones and proteasomes. *Mol. Cell. Biol.* **24**, 2682–2697 (2004).
- Karpova, T. S. *et al.* Concurrent fast and slow cycling of a transcriptional activator at an endogenous promoter. *Science* **319**, 466–469 (2008).
- MacArthur, S. *et al.* Developmental roles of 21 *Drosophila* transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biol.* **10**, R80 (2009).
- Dion, M. F. *et al.* Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**, 1405–1408 (2007).

- van Werven, F. J., van Teeffelen, H. A., Holstege, F. C. & Timmers, H. T. Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. *Nature Struct. Mol. Biol.* **16**, 1043–1048 (2009).
- Lieb, J. D., Liu, X., Botstein, D. & Brown, P. O. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nature Genet.* **28**, 327–334 (2001).
- Piña, B., Fernández-Larrea, J., García-Reyero, N. & Idrissi, F. Z. The different (sur)faces of Rap1p. *Mol. Genet. Genomics* **268**, 791–798 (2003).
- Buck, M. J. & Lieb, J. D. A chromatin-mediated mechanism for specification of conditional transcription factor targets. *Nature Genet.* **38**, 1446–1451 (2006).
- Holstege, F. C. *et al.* Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).
- Layer, J. H., Miller, S. G. & Weil, P. A. Direct transactivator-transcription factor IID (TFIID) contacts drive yeast ribosomal protein gene transcription. *J. Biol. Chem.* **285**, 15489–15499 (2010).
- Pelechano, V., Chavez, S. & Perez-Ortin, J. E. A complete set of nascent transcription rates for yeast genes. *PLoS ONE* **5**, e15442 (2010).
- Polach, K. J. & Widom, J. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254**, 130–149 (1995).
- Segal, E. & Widom, J. From DNA sequence to transcriptional behaviour: a quantitative approach. *Nature Rev. Genet.* **10**, 443–456 (2009).
- Lam, F. H., Steger, D. J. & O'Shea, E. K. Chromatin decouples promoter threshold from dynamic range. *Nature* **453**, 246–250 (2008).
- Pokholok, D. K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517–527 (2005).
- Anderson, J. D., Lowary, P. T. & Widom, J. Effects of histone acetylation on the equilibrium accessibility of nucleosomal DNA target sites. *J. Mol. Biol.* **307**, 977–985 (2001).
- John, S. *et al.* Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nature Genet.* **43**, 264–268 (2011).
- Zanton, S. J. & Pugh, B. F. Changes in genomewide occupancy of core transcriptional regulators during heat stress. *Proc. Natl Acad. Sci. USA* **101**, 16843–16848 (2004).
- Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366 (2009).
- Koerber, R. T., Rhee, H. S., Jiang, C. & Pugh, B. F. Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome. *Mol. Cell* **35**, 889–902 (2009).
- Ganapathi, M. *et al.* Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. *Nucleic Acids Res.* **39**, 2032–2044 (2011).
- Mukherjee, S. *et al.* Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nature Genet.* **36**, 1331–1339 (2004).
- Bosisio, D. *et al.* A hyper-dynamic equilibrium between promoter-bound and nucleoplasmic dimers controls NF- κ B-dependent gene activity. *EMBO J.* **25**, 798–810 (2006).
- Voss, T. C. *et al.* Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. *Cell* **146**, 544–554 (2011).
- McNally, J. G., Muller, W. G., Walker, D., Wolford, R. & Hager, G. L. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287**, 1262–1265 (2000).
- Perlmann, T., Eriksson, P. & Wrangé, O. Quantitative analysis of the glucocorticoid receptor-DNA interaction at the mouse mammary tumor virus glucocorticoid response element. *J. Biol. Chem.* **265**, 17222–17229 (1990).
- Gorski, S. A., Snyder, S. K., John, S., Grummt, I. & Misteli, T. Modulation of RNA polymerase assembly dynamics in transcriptional regulation. *Mol. Cell* **30**, 486–497 (2008).
- Rossetti, L. *et al.* Specific interactions of the telomeric protein Rap1p with nucleosomal binding sites. *J. Mol. Biol.* **306**, 903–913 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Kaplan and O. Rando for help with their turnover model, T. Palpant and S. Adar for help with time course experiments, and A. Leonardo Iniguez and H. Rosenbaum of Roche Nimblegen for pre-release custom HD4 12-plex microarrays. This work was supported by the US National Institutes of Health (NIH) Grant R01-GM072518 (to J.D.L.), and the intramural program of the NIH, National Cancer Institute, Center for Cancer Research (to J.G.M. and F.M.). F.M. was also supported in part by the Region Ile-de-France in the framework of C'Nano IdF, the nanoscience competence center of Paris Region.

Author Contributions C.R.L., S.E.H. and J.D.L. designed the study. C.R.L. and S.E.H. performed the experiments. F.M. developed and implemented the binding dynamics model. C.R.L., F.M., J.G.M. and J.D.L. performed data analysis. C.R.L., F.M., J.G.M. and J.D.L. wrote the paper.

Author Information Data has been deposited in the Gene Expression Omnibus under accession numbers GSE32351 (ChIP-on-chip data), GPL14612 (ChIP platform), GSM677030–GSM677033 (RNA expression array data) and GPL4414 (expression platform). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.D.L. (jlleb@bio.unc.edu).

METHODS

Strain construction. The *RAP1* gene and promoter was cloned into the pRS403 plasmid and integrated by homologous recombination into the *HIS3* locus of the BY4741 *S. cerevisiae* strain. The two copies of *RAP1* were then sequentially tagged using the 9× Myc epitope from pYM20:hphNT1 at the *HIS3* copy of *RAP1* and the 3× Flag tag from p3Flag-KanMX at the endogenous *RAP1* copy. The *HIS3* copy of the *RAP1* promoter was replaced using homologous recombination by amplifying the *GALL:natNT2* promoter from the pYM-N27 plasmid. Integrations were confirmed using PCR and western blot analysis. The *BARI* gene was knocked out by homologous recombination using a *LEU2* gene amplified from pRS405.

Time course. Yeast strains were grown overnight in YPD (yeast extract 1%, peptone 2% and dextrose 2%) and used to inoculate 800 ml of YPR (yeast extract 1%, peptone 2% and raffinose 2%) to an attenuation at 600 nm ($D_{600\text{ nm}}$) of 0.2 (Genesys 20 Spectrophotometer) in a 4-l Erlenmeyer flask. These cells were grown to an D_{600} of 0.4 and subsequently arrested using 5 μM alpha factor (400 μl of 10 mM, GenScript) until 95% of the yeast cells were unbudded (~3 h). Cells were then induced by adding 40% galactose to a final concentration of 2%. At this time, additional alpha factor was added (400 μl of 10 mM, GenScript). Samples were collected at time points 0, 10, 20, 30, 40, 50, 60, 90, 120 and 150 min after galactose induction. At each time point, 35 ml of culture was taken and added immediately to 37% formaldehyde to a final concentration of 1% for 20 min. Thirteen millilitres were taken for subsequent RNA preparation. Two millilitres were taken for protein preparation by pelleting cells and heating at 95 °C for 5 min in 0.06 M Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% 2-mercaptoethanol and 0.0025% bromophenol blue. All samples were frozen immediately in liquid nitrogen.

Turnover model. A mathematical model is required to interpret the data, and to obtain binding turnover rates. We used a modified version of a histone H3 turnover model⁶. The original H3 turnover model assumed that there was no competitor protein present before its induction⁶. We were also unable to detect the presence of the Rap1 competitor protein before induction using western blot analysis. Nevertheless, at each locus we consistently measured a non-zero competitor signal from the microarray even before the competitor was induced. This probably reflects the nonspecific background of our microarrays. Most of the steps that could contribute to this noise—for example, non-specific pull down from the beads, site-specific variations in the DNA amplification or nonspecific binding bias in hybridization—would affect the constitutive and competitor signal equally, and therefore we assume for simplicity that the total nonspecific background signal is approximately the same for the constitutive signal and for the competitor signal in our modified turnover model. We assume that at each binding site the measured immunoprecipitation signal ($mIP(t)$) is the true immunoprecipitation signal ($IP(t)$) plus the background ($BGD(t)$):

$$mIP(t) = IP(t) + BG(t) \quad (1)$$

We assume that at the beginning of the experiment (before induction) the true IP signal of the competitor is zero. The background signal at the start of the experiment is therefore the signal measured for the competitor protein A at time 0:

$$mIP_A(0) = BG(0) \quad (2)$$

The measured background signal will generally be time-dependent because our data showed that the measured raw intensities of the immunoprecipitation signals for the constitutive and competitor Rap1 proteins fluctuated from one time point to the next, even though their relative proportions remained roughly the same. This suggests that there are systematic variations in either the ChIP conditions or the microarray imaging conditions from one time point to the next, which would also probably influence the background signal.

The systematic changes in either the ChIP or imaging conditions can be quantified by comparing the total signal of constitutive Rap1 plus competitor Rap1 at each binding site as a function of time. We assume that the addition of competitor does not change total occupancy⁶ (Supplementary Figs 3 and 4). Thus, at each binding site, the ratio of the total signal (constitutive plus competitor Rap1) at time t versus time 0 generates a scaling factor to account for systematic fluctuations over time. This scaling factor (the part in brackets in equation (3)) can be used to calculate the background signal at time t based on the background at time 0:

$$BGD(t) = BG(0) \times \left(\frac{IP_A(t) + IP_B(t)}{IP_A(0) + IP_B(0)} \right) \quad (3)$$

With this formula, we can calculate an occupancy ratio in the presence of background signal. First note that the occupancy ratio $R(t)$ in the absence of background is defined as the ratio of the immunoprecipitates of the competitor and constitutive signals:

$$R(t) = IP_A(t)/IP_B(t) \quad (4)$$

We define a measured occupancy ratio $mR(t)$ that includes the background signal:

$$mR(t) = mIP_A(t)/mIP_B(t) = [IP_A(t) + BG(t)]/[IP_B(t) + BG(t)] \quad (5)$$

where the second equality arises by substitution from equation (1), assuming that the background is the same in the competitor and constitutive signals. Using equations (2), (3) and (4), equation (5) can be rewritten as:

$$mR(t) = [R(t) + C_0(1 + R(t))]/[1 + C_0(1 + R(t))] \quad (6)$$

where $C_0 = mIP_A(0)/(IP_A(0) + IP_B(0))$. This constant can be expressed in terms of measurable quantities by using equations (1), (2) and the previously stated assumption $IP_A(0) = 0$ to yield:

$$C_0 = \frac{mR(0)}{1 - mR(0)} \quad (7)$$

where $mR(0)$ is the measured occupancy ratio at time 0. In practice, we calculated C_0 by averaging over the first three time points, and found that they all showed no detectable competitor signal. With this estimate of C_0 , equation (6) enables an occupancy ratio to be calculated in the presence of a microarray background signal by using the occupancy ratio R_t calculated in the absence of background⁶.

$R(t)$ is the probability that a locus is occupied by the competitor protein divided by the probability that it is occupied by the constitutive protein⁶. If P is the probability that the competitor occupies a given locus, then the probability that the constitutive protein occupies the locus is $1 - P(t)$, and so $R(t)$ becomes:

$$R(t) = \frac{P(t)}{1 - P(t)} \quad (8)$$

This probability satisfies the following differential equation⁶:

$$\frac{d}{dt} P(t) = \lambda \left(\frac{A(t)}{A(t) + B(t)} - P(t) \right) \quad (9)$$

Here λ is the turnover rate at each locus, and $A(t)$ and $B(t)$ are the cellular concentrations of the free competitor and constitutive proteins. We measured $A(t)$ and $B(t)$ at all time points using western blot analysis. To determine the turnover rate λ for each locus we tuned λ to fit the measured occupancy ratio $mR(t)$ at that locus. Specifically, we varied λ in equation (9) such that the value of $R(t)$ obtained from equation (8) yields the best fit to our measured occupancy ratio when $R(t)$ is substituted into equation (6).

The modified turnover model (equation (6)) was implemented in Matlab 2009b (The MathWorks) and equation (9) was solved numerically using the *ODE45* function. The Matlab routine 'lsqcurvefit' was used to fit the models to experimental data and extract the turnover rate λ . We sampled a range of different starting guesses to avoid the detection of local minima. The Matlab source code for the modified turnover model is available online (<http://code.google.com/p/cc-process/>).

Plasmids. The following plasmids were used in generation of the Rap1 turnover strain: pRS403 (ref. 31), pRS405 (ref. 31), pYM20:hphNT1 (ref. 32), p3FLAG-KanMX (ref. 33) and pYM-N27 (ref. 32).

Chromatin immunoprecipitation and DNA amplification. Chromatin immunoprecipitation was performed on whole-cell extract from crosslinked cells as described previously, using anti-Flag (M2, Sigma), anti-Myc (clone 9E10, Millipore), and anti-Rap1 (y-300, Santa Cruz Biotechnology)⁸. Immunoprecipitated and/or input DNA was amplified using the GenomePlex Complete Whole-Genome Amplification (WGA) kit (WGA2-50RXN, Sigma) and then re-amplified using GenomePlex WGA Reamplification Kit (WGA3-50RXN, Sigma) using the manufacturer's protocols. DNA was purified using Zymo columns according to the manufacturer's instructions (Zymo Research).

Hybridization and processing of data from high-resolution HD4 microarrays. For Nimblegen high-resolution HD4 microarrays, amplified ChIP material was sent directly to Nimblegen where it was labelled and hybridized according to protocols in chapters 3 and 4 of the NimbleGen Arrays User's Guide ChIP-chip Analysis, Version 3.1, 27 May 2008. Bi-weight mean scaled ratios are used as input for lowess normalization. All HD4 array data are deposited in the Gene Expression Omnibus (GEO) under accession GSE32351.

Modified lowess normalization. Standard lowess normalization results in depressed binding ratios at the most highly enriched probes in ChIP-chip experiments. We therefore implemented a modified lowess normalization designed specifically for ChIP-chip based on the method described previously⁷. Our method varied from the previously published method⁷ in that we defined the 'enriched' group of probes based on the sites that we used to define Rap1 target enrichment for our turnover time course. We excluded all probes within $\pm 2,000$ bp of a Rap1 binding site, and used all remaining probes as the reference group to calculate the lowess function for normalization (Supplementary Fig. 8a–d). Each time point is normalized separately but we use the same group of reference probes for the normalization. Although we believe that using this modified lowess normalization

approach is the most appropriate way to normalize the data, we find qualitative and quantitatively similar Rap1 turnover values without normalization (data not shown).

Hybridization and processing of data from low-resolution PCR-based arrays. One microgram of amplified DNA was labelled with either 2'-deoxyuridine, 5'-triphosphate (dUTP) Cy5 (PA55022, GE Healthcare) or Cy3 (PA53022, GE Healthcare) for low-resolution PCR-based arrays. Purified labelled DNA was hybridized to PCR-based arrays representing the whole yeast genome and covering all coding and non-coding regions at an average resolution of approximately 800 bp (ref. 10). The time course was performed in duplicate, one in each dye orientation, with the Myc and Flag samples then comparatively hybridized to an array for each time point. Arrays were scanned using an Axon 4000B scanner and analysed using Genepix 6.0 software (Axon). Only spots with <10% saturated input pixels and a signal intensity of greater than 500 (background-corrected sum of medians for both channels) were used for the analysis. Data were further normalized in the UNC microarray database with the normalized median \log_2 ratio of Rap1-Myc/Rap1-Flag being used for further analysis. All low-resolution array data are deposited in GEO under accession GSE27377. We did not use these ChIP-chip data in any of our analyses except in Supplementary Fig. 3.

Reverse transcription, complementary-DNA labelling and expression arrays. Total RNA was extracted by the hot phenol method as previously described³⁴. Total RNA (30 μ g) was reverse transcribed into cDNA using reagents and protocols provided with SuperScript II reverse transcriptase (Invitrogen; Cat. No. 18064-014) containing an amino-allyl-dUTP mix (50 \times aa-dUTP mixture; 1 mg amino-allyl dUTP (Sigma)) dissolved with 32 μ l of 100-mM dATP, dGTP, dCTP, 12.7 μ l of 100-mM dTTP, and 19.3 μ l of dH₂O, and an anchored oligo dT primer (22-mer; IDT). Reactions were incubated for 2 h at 42 °C, then heated at 95 °C for 5 min and snap cooled on ice. RNA was hydrolysed by addition of 13 μ l of 1-N NaOH and 1 μ l of 0.5-M EDTA followed by incubation at 67 °C. Reactions were then neutralized with 50 μ l of 1 M HEPES pH 7.5. cDNA was purified on Zymo columns (Zymo Research; D4003) using a seven-volume excess of DNA binding buffer. cDNA was eluted from columns using 5 μ l of 50-mM sodium bicarbonate pH 9.0. cDNA was fluorescently labelled using Amersham CyDye Post-Labeling Reactive Dye Packs (RPN5661). Each dye pack was resuspended in 11 μ l DMSO and 3 μ l of mixture was used per reaction. Cy dyes and aa-dUTP cDNAs were allowed to couple for 2 h in the dark. Labelled cDNAs were cleaned up using Zymo columns with a seven-volume excess of DNA binding buffer and eluted with 10 mM Tris-Cl pH 8.0 and hybridized to arrays as described previously.

For comparative hybridization, input genomic DNA from the experimental Rap1 turnover strain was extracted using phenol chloroform. Four micrograms of genomic DNA was denatured at 100 °C with 10 μ g of random hexamer (IDT) then snap cooled on ice for 10 min. Samples were then incubated with 50 units of Klenow (exo-) (New England Biolabs (NEB)) and 1 \times Buffer 2 (NEB) in a total volume of 50 μ l at 37 °C for 2 h. Samples were cleaned up with Zymo columns, eluted in 5 μ l of 5-mM sodium bicarbonate pH 9.0 and coupled to Cy dyes as for cDNA. Expression studies were performed on PCR-based arrays that were prepared, processed and analysed as for the low-resolution ChIP arrays¹⁰.

Defining regions of Rap1 enrichment. Rap1 ChIP-seq data from yeast strain BY4741 grown in YPD (yeast extract 1%, peptone 2% and dextrose 2%) were used to determine precise sites of Rap1 binding. Peaks and peak summits were identified using model-based analysis for ChIP-Seq (MACS) with a bandwidth of 300 and a *P*-value cutoff of 1×10^{-5} . Peaks in our turnover data set were identified on total Rap1 occupancy at time 0 using Peakpicker³⁵ to ensure that we identified all Rap1 peaks that were present in our turnover conditions. For analysis, we then used only MACS ChIP-seq peak regions that had at least 1 bp of overlap with our time-course peaks, and a *z* score of >1.5 at time 0. Seven regions with a *z* score of >1.5 at time 0 that were identified at time 0 of the Rap1 time course but not of the ChIP-seq experiment were also included to ensure full representation of Rap1-enriched regions in our experiment. Of the 457 total Rap1 peak regions identified, 18 were not analysed. Fifteen targets had an estimated residence time of under 500 s, which is too short to measure with our system (Supplementary Fig. 6). Three targets that had residence times that exceeded 1×10^4 s and showed exceptionally poor fits to the model were also excluded. The average \log_2 Myc/Flag level for all probes which fell within ± 150 bp of peak summits were averaged to generate a

Myc/Flag value for each time point for each target. On average, eight probes contributed to the Myc/Flag signal for Rap1 targets. Peak summits were used to assign target regions to promoters or coding regions for further analysis.

Telomeric regions were tiled using only uniquely mapping probes, making signal discontinuous in these regions and making peak calling difficult. For this reason, telomeres were defined by annotations from the *Saccharomyces* genome database (<http://www.yeastgenome.org/>). We excluded telomeres from any analysis that relied on our turnover metric because they contain many arrayed Rap1 binding sites within their AC-rich repeats. In theory, as the number of Rap1-binding sites detected by an individual microarray probe increases, the probability that either isoform of Rap1 will be detected at that probe increases. This violates some assumptions of our turnover metric, which would theoretically lead to artificially short residence-time estimates. Despite this, empirically we see no relation between Rap1 residence times and motif number or density (Supplementary Fig. 10).

Motif discovery. The 439 Rap1-bound target regions (excluding telomeres) were placed into four categories based on their turnover properties: longest (110 targets), long (110 targets), short (110 targets) and shortest (109 targets). The DNA sequences for each Rap1 target region in each group were then used as input for the web-based interface for BioProspector³⁶ (<http://ai.stanford.edu/~xslu/BioProspector/>). Default parameters were with the exception of the width of the first motif block, which was changed to '13' and '*S. cerevisiae* intergenic' was used as a genome-background model. Rap1's telomeric motif was determined from the full telomeric sequences of the 26 telomeres that were uniquely mappable on our arrays. Weblogo³⁷ (<http://weblogo.berkeley.edu/logo.cgi>) was used to generate a visual representation of the position-weight matrix output from BioProspector. The 439 Rap1 targets were similarly grouped by their occupancy properties to determine Rap1 motifs for Rap1 targets grouped by occupancy. The default settings on the motif scanning program Clover³⁸ were used to detect Rap1 motifs genome-wide using a previously published Rap1 position-weight matrix¹⁰.

External data sets. Values from existing data sets with a one-to-one correspondence to the arrayed elements in our study were used as published. For data sets derived from arrays that did not match our probe set, \log_2 ratios and *z* scores were calculated for each array probe, for each replicate of the external data set. *Z* scores were defined as the number of standard deviations that a probe's \log_2 ratio was from the mean \log_2 ratio of all probes on the array. In cases with several replicates, average *z* scores were used to represent each probe. To map the data back to our experiments, the average *z* scores of the array probes for the specific data set that were contained within the promoter or coding region assigned to each Rap1 target were used for comparison. For histone H3 turnover data, the highest value for a probe that fell within promoters associated with peak summits for target regions was used for our analysis⁶. For Rap1 nucleosome interaction data we summed all the detected interactions that fell within each Rap1 target region.

- Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
- Janke, C. et al. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).
- Gelbart, M. E., Rechsteiner, T., Richmond, T. J. & Tsukiyama, T. Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Mol. Cell. Biol.* **21**, 2098–2106 (2001).
- Hoffman, C. S. in *Current Protocols in Molecular Biology*, Vol. 2 (eds Ausubel, F. M. et al.) 13.11.1–13.11.4 (John Wiley and Sons, 1997).
- Cesarini, M., Cittaro, D., Brofzi, A., Pelicci, P. G. & Luzi, L. CARPET: a web-based package for the analysis of ChIP-chip and expression tiling data. *Bioinformatics* **24**, 2918–2920 (2008).
- Liu, X., Brutlag, D. L. & Liu, J. S. BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac. Symp. Biocomput.* **6**, 127–138 (2001).
- Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
- Frith, M. C. et al. Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res.* **32**, 1372–1381 (2004).

A new understanding of the decoding principle on the ribosome

Natalia Demeshkina^{1*}, Lasse Jenner^{1*}, Eric Westhof², Marat Yusupov¹ & Gulnara Yusupova¹

During protein synthesis, the ribosome accurately selects transfer RNAs (tRNAs) in accordance with the messenger RNA (mRNA) triplet in the decoding centre. tRNA selection is initiated by elongation factor Tu, which delivers tRNA to the aminoacyl tRNA-binding site (A site) and hydrolyses GTP upon establishing codon–anticodon interactions in the decoding centre^{1–9}. At the following proofreading step the ribosome re-examines the tRNA and rejects it if it does not match the A codon^{2,3,10–14}. It was suggested that universally conserved G530, A1492 and A1493 of 16S ribosomal RNA, critical for tRNA binding in the A site^{15–17}, actively monitor cognate tRNA¹⁸, and that recognition of the correct codon–anticodon duplex induces an overall ribosome conformational change (domain closure)¹⁹. Here we propose an integrated mechanism for decoding based on six X-ray structures of the 70S ribosome determined at 3.1–3.4 Å resolution, modelling cognate or near-cognate states of the decoding centre at the proofreading step. We show that the 30S subunit undergoes an identical domain closure upon binding of either cognate or near-cognate tRNA. This conformational change of the 30S subunit forms a decoding centre that constrains the mRNA in such a way that the first two nucleotides of the A codon are limited to form Watson–Crick base pairs. When U•G and G•U mismatches, generally considered to form wobble base pairs, are at the first or second codon–anticodon position, the decoding centre forces this pair to adopt the geometry close to that of a canonical C•G pair. This by itself, or with distortions in the codon–anticodon mini-helix and the anticodon loop, causes the near-cognate tRNA to dissociate from the ribosome.

We determined six X-ray structures of the 70S ribosome at 3.1–3.4 Å resolution (Supplementary Tables 1 and 2) programmed by 30-nucleotide-long mRNAs with the AUG codon and tRNA^{fMet} in the peptidyl tRNA-binding site (P site) and the A site occupied by tRNA_{2^{Leu}} or tRNA^{Tyr} (Fig. 1a and Methods). In one set of experiments, tRNA_{2^{Leu}} and tRNA^{Tyr} were bound to their respective cognate codons CUC and UAC in the A site. In a second set of experiments, we modelled near-cognate states of the ribosome (Supplementary Fig. 1). These states of the ribosome naturally occur during protein synthesis but with low probability because binding of cognate tRNA is kinetically

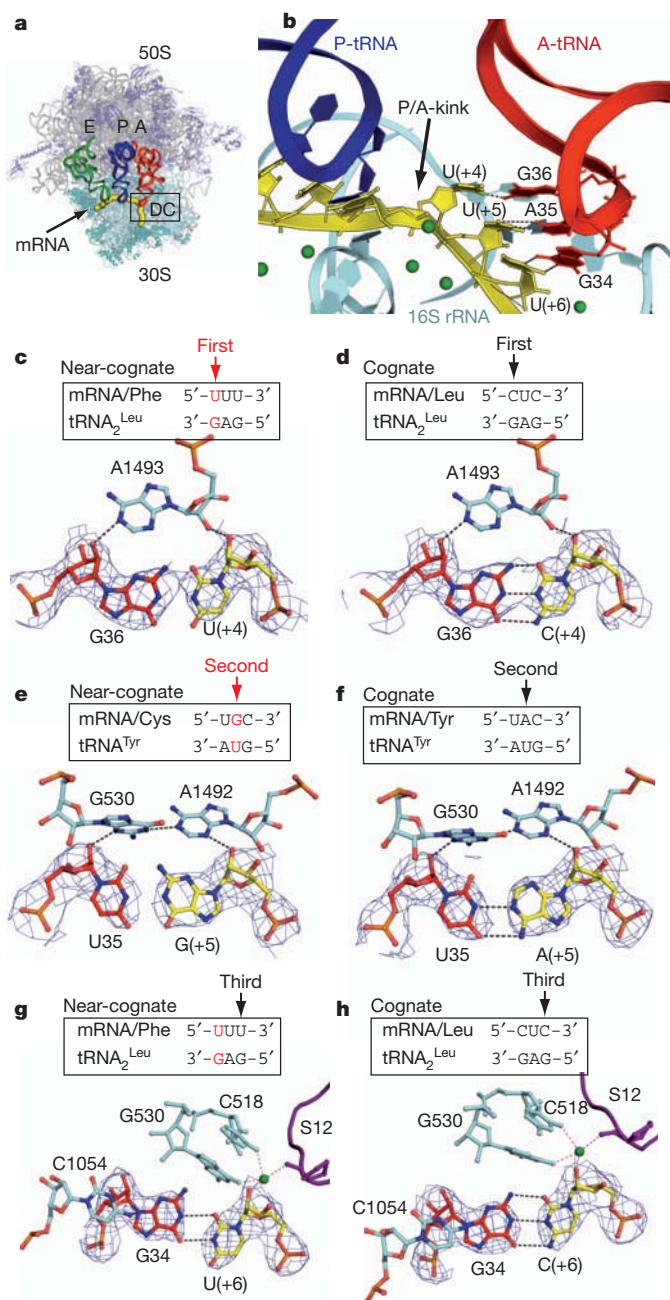


Figure 1 | Codon–anticodon interactions in the decoding centre on the 70S ribosome. **a**, The mRNA path on the 70S ribosome with the decoding (DC) area indicated. **b**, Close-up view of the mRNA P/A kink with near-cognate tRNA_{2^{Leu}}. Magnesium ions are in green. **c**, **d**, The first base pairs of the near-cognate (**c**) and cognate (**d**) codon–anticodon duplexes and their interactions with A1493 of 16S rRNA. **e**, **f**, The second base pairs of the near-cognate (**e**) and cognate (**f**) codon–anticodon duplexes and their interactions with G530 and A1492 of 16S rRNA. **g**, **h**, A classical wobble U•G pair (**g**) versus canonical C•G interactions (**h**); a magnesium ion interacting with the base pair is coordinated by protein S12 and part of 16S rRNA. All graphical representations were prepared with PyMol. Probable hydrogen bonds within 3 Å distance are indicated by dashed lines; $2F_o - F_c$ electron density maps are contoured at 1.2 sigma.

¹Département de Biologie et de Génétique Structurale, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch 67400, France; CNRS, UMR7104, Illkirch 67400, France; INSERM, U964, Illkirch 67400, France; Université de Strasbourg, Strasbourg 67000, France. ²Architecture et Réactivité de l'ARN, Université de Strasbourg, Institut de Biologie Moléculaire et Cellulaire, CNRS, Strasbourg 67084, France.

*These authors contributed equally to this work.

favoured. We made the ribosome accept near-cognate tRNA by giving it only one type of tRNA carrying a mismatch to the A codon along with tRNA^{Met} for the P site. In these complexes the A site was filled either by tRNA₂^{Leu} and codon UUU with a U•G mismatch in the first position of the codon–anticodon helix or by tRNA^{Tyr} and codon UGC with a G•U mismatch in the second position.

As shown earlier, the mRNA forms a kink between the P and A codons (P/A kink), a universal feature of the mRNA path on the 70S ribosome that is stabilized by the P site tRNA, 16S ribosomal RNA (rRNA) and magnesium ions^{7,20,21} (Fig. 1b and Supplementary Fig. 2). The single mismatch states described above represent bona fide near-cognate complexes expected to have standard wobble U•G and G•U base pairs. At our data resolution (3.1–3.5 Å) we can confidently assign the general base pairing (Supplementary Figs 3 and 4). The electron density maps unambiguously demonstrate that U4 and G5 of the A codons UUU and UGC do not show the anticipated wobble interactions with G36 in tRNA₂^{Leu} and U35 in tRNA^{Tyr}, respectively. Instead, U4•G36 and G5•U35 at the first and second positions of

the codon–anticodon duplexes form base pairs similar to a standard Watson–Crick G•C pair (Fig. 1c, e and Supplementary Fig. 5). G•C-like G•U or G•T pairs have been shown earlier for RNA and DNA in other X-ray structures^{22,23}. When U•G is at the third codon–anticodon position we observe standard U•G wobble pairing (Fig. 1g, h). Unexpectedly, nucleotides A1493, A1492 and G530 of 16S rRNA in helix 44 (h44), which contact the first and the second pairs of the codon–anticodon helix, interact with these unusual U4•G36 and G5•U35 pairs identically to the way they interact with canonical Watson–Crick base pairs C4•G36 and A5•U35 (Fig. 1c–f). These findings are in contradiction with studies where these nucleotides were given a role as monitors and discriminators of canonical Watson–Crick pairs in the decoding process^{18,19}. Our structures show that G530, A1492 and A1493 form a static part of the decoding centre, defining its spatial and stereochemical properties (Fig. 2a, b).

The observed non-wobble U•G differs from previous X-ray studies that were based on the 30S subunit alone¹⁹ (Fig. 2c and Supplementary Figs 5 and 6). For example, for the study of the mismatch at the first

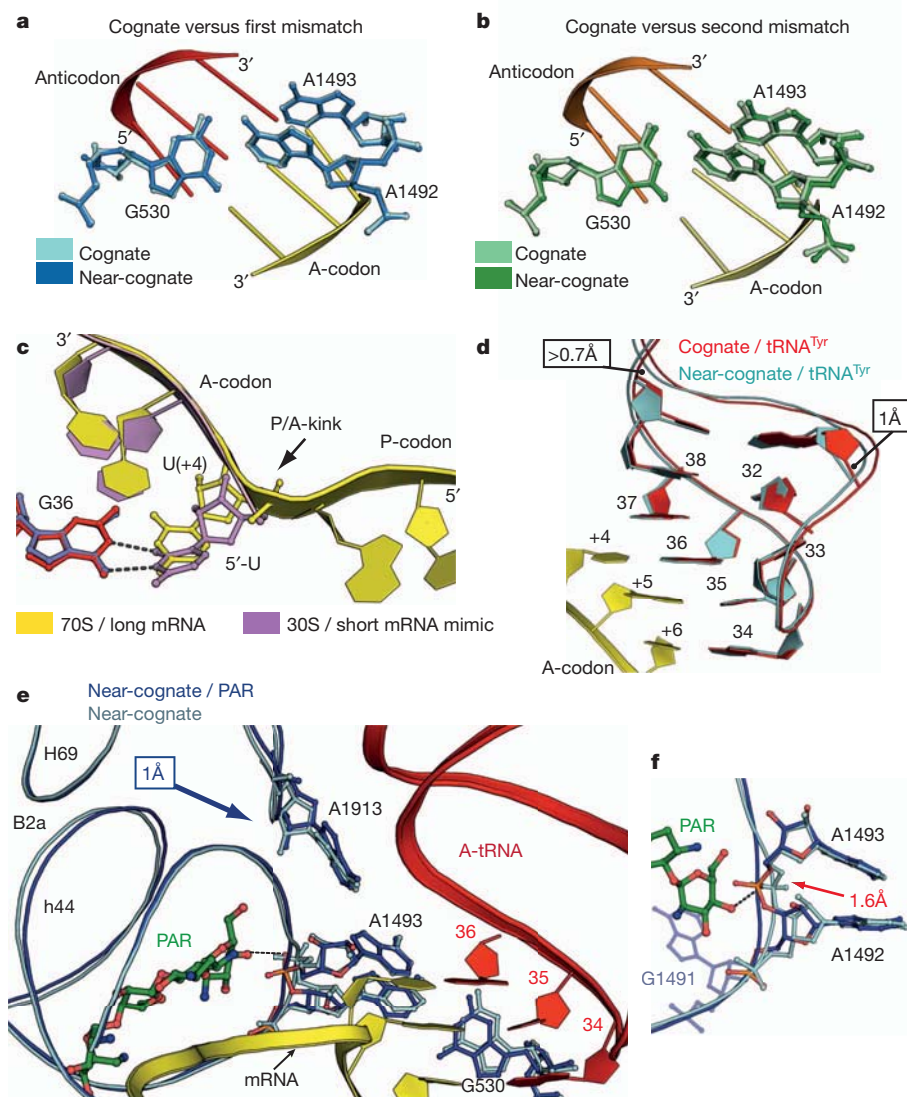


Figure 2 | The nature of the decoding centre. **a, b**, The overall conformations of universally conserved G530, A1492 and A1493 of 16S rRNA in the cognate structures are identical to those in the near-cognate models when the mismatches are at the first (**a**) or second (**b**) codon–anticodon positions. **c**, Differences between the position of the first uridine in the UUU codon base-paired to the GAG anticodon of tRNA₂^{Leu} from our 70S structure and from the 30S model¹⁹. **d**, Comparison of the anticodon loops of tRNA^{Tyr} in the cognate

(red) and near-cognate (cyan) states. **e**, Rearrangements of rRNA helices h44 and H69 in the near-cognate state upon binding of the aminoglycoside paromomycin (PAR). The near-cognate structures with tRNA^{Tyr} are shown (a similar effect of PAR is observed with tRNA₂^{Leu}, see also Supplementary Fig. 10). **f**, Magnified view of the changes in the A1493 phosphate position induced by PAR. Superimpositions in **d**, **e** and **f** were performed using 23S rRNA as reference.

position, 30S crystals soaked with an anticodon stem-loop of tRNA^{Leu}₂ and a hexauridine (U₆) mRNA displayed a classical wobble U•G pair. There the P codon was mimicked by the 3'-end of 16S rRNA, so the U₆ mRNA could only bind to the A site and downstream, leading to a situation where the mRNA was not covalently linked between the P and A codons. By superimposing the A site GAG anticodons from our near-cognate tRNA^{Leu}₂ structure and the 30S model we found that the first nucleotide of the A codon is positioned differently (Fig. 2c and Supplementary Fig. 6). Because it does not have the natural restraint coming from being covalently bound to the P codon, this first nucleotide in the A codon has the freedom to move so it can form a wobble U•G pair. However, in our structure the P/A mRNA kink specifically directs the first nucleotide of the A codon to form Watson–Crick like interactions with G36 of tRNA^{Leu}₂ (Fig. 1c).

The positional restrictions imposed by the decoding centre on the first two near-cognate U•G codon–anticodon pairs may result in differences in geometry compared with the corresponding cognate helices (Supplementary Figs 7 and 8). Although, the resolution of the data sets does not allow us to determine the exact value for these deviations in the base pairs, the tendencies are clear. A noticeable change is an increase in buckling at the first and third codon–anticodon positions with the U4•G36 mismatch (Supplementary Fig. 7) and at the second and third positions with the G5•U35 mismatch (Supplementary Fig. 8). These deviations could disturb the base stacking network within the near-cognate codon–anticodon helices and deform the entire anticodon loop structure²⁴ (Fig. 2d, Supplementary Fig. 9 and Supplementary Movie 1). This deformation might influence the position of helix 69 (H69) of 23S rRNA, whose universally conserved A1913 protrudes into the decoding centre^{20,21} (Fig. 2e).

Additional structures of the near-cognate states determined in the presence of the miscoding aminoglycoside paromomycin (Supplementary Tables 1 and 2) reveal a movement of H69 accompanied by

rearrangements of the intersubunit bridge B2a between h44 and H69 (Fig. 2e and Supplementary Movie 2). These distortions are most probably caused by binding of the antibiotic that strongly shifts the A1493 phosphate group (Fig. 2f). Although this shift does not alter much the interactions of A1493 with U4•G36 and A1492/G530 with G5•U35 (Supplementary Figs 7g and 8h), these local changes modulate the B2a bridge. H69 is displaced towards the tRNA, which probably enhances the interaction surface of H69 with the tRNA D-stem. In the presence of paromomycin the position of H69 is closer to that observed for the cognate state (Supplementary Fig. 10). Furthermore, displacement of the A1493 phosphate group relaxes the decoding pocket from the side of the A codon (Fig. 2e) and changes the deformation of the near-cognate codon–anticodon helix (Supplementary Movie 2). This novel understanding of the paromomycin action therefore differs from the previously suggested mechanism in decoding where it was proposed to influence the monitoring capabilities of A1492 and A1493 (ref. 18). The observed moderate structural rearrangements with paromomycin are consistent with its measured effect at the proofreading step²⁵.

We find that both near-cognate tRNAs induce rearrangements of the 30S subunit known as domain closure¹⁹ (that is, shoulder movement and head rotation) as described for cognate tRNA²¹ (Fig. 3a). This implies that domain closure is an inherent quality of the ribosome in response to binding of any tRNA to the A site¹² and is prerequisite for formation of the decoding pocket.

Initially, the mechanism underlying the decoding process was deduced from pioneer X-ray structures of the isolated 30S subunit where crystals were soaked with U₆ and anticodon stem-loop mimicking mRNA and tRNA^{18,19}. Besides those limitations, all attempts to model near-cognate states on the 30S subunit were performed in the presence of paromomycin, which, on the one hand, stimulated an ordered binding of the near-cognate tRNA analogs, but, on the other,

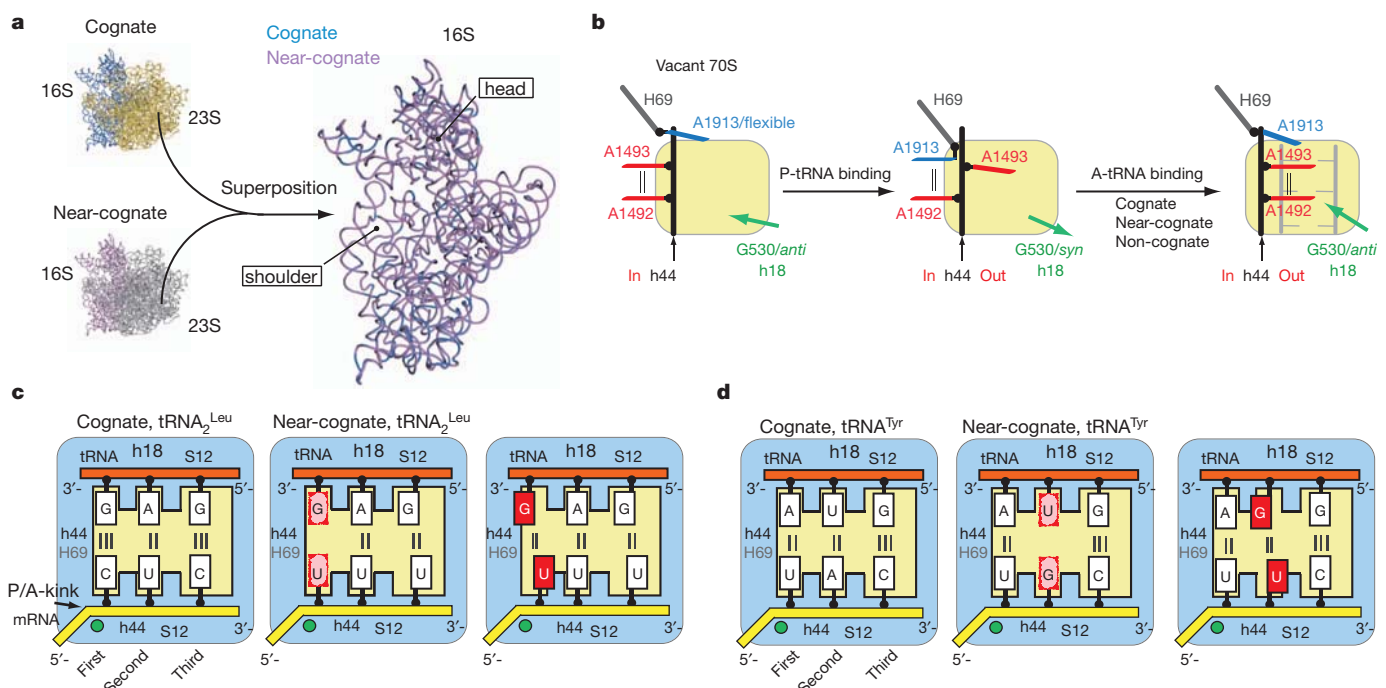


Figure 3 | The principle of decoding. **a**, Superposition of 23S rRNA from near-cognate and cognate structures with tRNA^{Leu}₂ shows identical domain closure in the near-cognate and cognate states. **b**, Representation of conformations of the core nucleotides composing the decoding centre depending on its functional states (data with the vacant 70S ribosome are unpublished). **c**, **d**, Illustration of the decoding principle: together with the constraints imposed on the A codon by the P/A kink coordinated by a magnesium ion (green sphere), the DC (h18, h44 and protein S12 from the

small subunit and H69 from the large subunit) restricts the allowed geometry of the first two nucleotides of the codon. No such restraints are imposed on the third base pair. A near-cognate tRNA with G•U in the first or second position is forced to form Watson–Crick-like base pairs (middle panels). This creates repulsion or requires energy for tautomerization (shown in pink), which by itself can be the source of the tRNA discrimination. The right panels illustrate the impossible situation when standard wobble base pairs (shown in red) occur in either the first or second positions of the codon–anticodon duplexes.

made it difficult to distinguish the independent effect of the mismatch on the 30S decoding centre¹⁹. The cognate and near-cognate models of the 70S ribosome described here, with or without paromomycin, together with our previous structures^{12,21}, give rise to novel insights into decoding with revisited roles for universally conserved nucleotides G530, A1492 and A1493 of 16S rRNA (Fig. 3b). Although these nucleotides form extensive contacts to the A-minor groove of a codon–anticodon helix, G530, A1492 and A1493 do not actively sense the correct Watson–Crick base-pairing geometry and thus do not discriminate against near-cognate tRNA.

We propose that upon binding of cognate or near-cognate tRNA to the 70S ribosome, the small subunit undergoes domain closure around the anticodon loop of the tRNA. The closure results in formation of a tight decoding centre that restricts the first two nucleotides of the A codon to form exclusively Watson–Crick base pairs with the tRNA anticodon (Fig. 3c, d). Owing to our current data resolution, we cannot precisely identify the hydrogen bond pattern of the mismatches in the near-cognate codon–anticodon helices, but tautomerism is a plausible chemical mechanism. An alternative explanation for the tRNA discrimination source could be repulsion in the U•G pair (Fig. 1c, e). Energy expenditure for formation of tautomers (or repulsion energy) could constitute the sole cause for the very efficient rejection of near-cognate tRNAs by the ribosome^{3,4}. Additionally, the observed deformation of the anticodon may lead to alterations in the codon–anticodon mini-helix and propagate through the rest of the near-cognate tRNA molecule, destabilizing it and causing the tRNA to dissociate from the ribosome (Supplementary Fig. 11). This corroborates the idea that evolutionarily tuned sequences of tRNAs play an active role in the tRNA selection process^{26–28}. Recent X-ray structures of the 70S ribosome with cognate tRNA and elongation factor Tu^{7,28} demonstrated that binding of an anticodon loop of tRNA in the decoding centre is nearly identical to that shown for accommodated tRNA. This information prompts us to hypothesize that the same mechanism described here for the proofreading step governs the initial tRNA selection step.

METHODS SUMMARY

Ribosomes were purified from *Thermus thermophilus* cells as described before²¹. For all complexes, mRNA, tRNA^{Met} and tRNA^{Leu} (or tRNA^{Tyr}) were present in fivefold stoichiometric excess of the ribosome concentration. Complexes with paromomycin were obtained by including the antibiotic (30 µM) into the incubation mixture. Crystals were grown at 24 °C by sitting-drop vapour diffusion as described before²⁰. All crystals belonged to space group *P*2₁2₁2₁ and contained two molecules per asymmetric unit. A very low dose mode with high redundancy was used for data collection²⁹. The structure²¹, with tRNAs, mRNA and metal ions removed, was used for refinement with Phenix³⁰. Throughout refinement, no base-pair restraints were used between tRNAs and mRNAs to avoid bias towards standard base-pair geometry.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 August 2011; accepted 1 February 2012.

Published online 21 March 2012.

- Rodnina, M. V. & Wintermeyer, W. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu. Rev. Biochem.* **70**, 415–435 (2001).
- Rodnina, M. V., Gromadski, K. B., Kothe, U. & Wieden, H. J. Recognition and selection of tRNA in translation. *FEBS Lett.* **579**, 938–942 (2005).
- Zaher, H. S. & Green, R. Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**, 746–762 (2009).
- Gromadski, K. B., Daviter, T. & Rodnina, M. V. A uniform response to mismatches in codon–anticodon complexes ensures ribosomal fidelity. *Mol. Cell* **21**, 369–377 (2006).
- Lee, T. H., Blanchard, S. C., Kim, H. D., Puglisi, J. D. & Chu, S. The role of fluctuations in tRNA selection by the ribosome. *Proc. Natl Acad. Sci. USA* **104**, 13661–13665 (2007).

- Li, W. *et al.* Recognition of aminoacyl-tRNA: a common molecular mechanism revealed by cryo-EM. *EMBO J.* **27**, 3322–3331 (2008).
- Schmeing, T. M. *et al.* The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688–694 (2009).
- Schuetz, J. C. *et al.* GTPase activation of elongation factor EF-Tu by the ribosome during decoding. *EMBO J.* **28**, 755–765 (2009).
- Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. The mechanism for activation of GTP hydrolysis on the ribosome. *Science* **330**, 835–838 (2010).
- Ehrenberg, M., Kurland, C. G. & Ruusala, T. Counting cycles of EF-Tu to measure proofreading in translation. *Biochimie* **68**, 261–273 (1986).
- Voorhees, R. M., Weixlbaumer, A., Loakes, D., Kelley, A. C. & Ramakrishnan, V. Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nature Struct. Mol. Biol.* **16**, 528–533 (2009).
- Jenner, L., Demeshkina, N., Yusupova, G. & Yusupov, M. Structural rearrangements of the ribosome at the tRNA proofreading step. *Nature Struct. Mol. Biol.* **17**, 1072–1078 (2010).
- Geggier, P. *et al.* Conformational sampling of aminoacyl-tRNA during selection on the bacterial ribosome. *J. Mol. Biol.* **399**, 576–595 (2010).
- Whitford, P. C. *et al.* Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. *RNA* **16**, 1196–1204 (2010).
- Moazed, D. & Noller, H. F. Binding of tRNA to the ribosomal A and P sites protects two distinct sets of nucleotides in 16S rRNA. *J. Mol. Biol.* **211**, 135–145 (1990).
- Powers, T. & Noller, H. F. Selective perturbation of G530 of 16S rRNA by translational miscoding agents and a streptomycin-dependence mutation in protein S12. *J. Mol. Biol.* **235**, 156–172 (1994).
- Yoshizawa, S., Fourmy, D. & Puglisi, J. D. Recognition of the codon–anticodon helix by ribosomal RNA. *Science* **285**, 1722–1725 (1999).
- Ogle, J. M. *et al.* Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292**, 897–902 (2001).
- Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721–732 (2002).
- Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).
- Jenner, L. B., Demeshkina, N., Yusupova, G. & Yusupov, M. Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nature Struct. Mol. Biol.* **17**, 555–560 (2010).
- BPS Database of RNA Base-pair Structures. <http://bps.rutgers.edu/bps> (2008).
- Bebenek, K., Pedersen, L. C. & Kunkel, T. A. Replication infidelity via a mismatch with Watson–Crick geometry. *Proc. Natl Acad. Sci. USA* **108**, 1862–1867 (2011).
- Auffinger, P. & Westhof, E. An extended structural signature for the tRNA anticodon loop. *RNA* **7**, 334–341 (2001).
- Pape, T., Wintermeyer, W. & Rodnina, M. V. Conformational switch in the decoding region of 16S rRNA during aminoacyl-tRNA selection on the ribosome. *Nature Struct. Biol.* **7**, 104–107 (2000).
- Cochella, L. & Green, R. An active role for tRNA in decoding beyond codon:anticodon pairing. *Science* **308**, 1178–1180 (2005).
- Dale, T. & Uhlenbeck, O. C. Amino acid specificity in translation. *Trends Biochem. Sci.* **30**, 659–665 (2005).
- Schmeing, T. M., Voorhees, R. M., Kelley, A. C. & Ramakrishnan, V. How mutations in tRNA distant from the anticodon affect the fidelity of decoding. *Nature Struct. Mol. Biol.* **18**, 432–436 (2011).
- Mueller, M., Wang, M. & Schulze-Briese, C. Optimal fine ϕ -slicing for single-photon-counting pixel detectors. *Acta Crystallogr. D* **68**, 42–56 (2012).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to C. Schulze-Briese and the staff at the Swiss Light Source (Switzerland) for help during synchrotron X-ray data collection. We thank S. Duclaud for ribosome preparation and the staff of the Structural Biology Department core facility at Institut de Génétique et de Biologie Moléculaire et Cellulaire, Université de Strasbourg. This work was supported by ANR BLAN07-3_190451 (to M.Y.), ANR-07-PCVI-0015-01 (to G.Y.), Fondation pour la Recherche Médicale en France (to N.D.) and by the European Commission SPINE2.

Author Contributions N.D. and L.J. conducted experiments and performed analysis. All authors discussed the results and commented on the manuscript.

Author Information The atomic coordinates and structure factors for the determined crystal structures are deposited in the Protein Data Bank under accession numbers 3TVF and 3TVE (cognate tRNA^{Leu} complex), 3UYE and 3UYD (near-cognate tRNA^{Leu} complex), 3UZ3 and 3UZ1 (near-cognate tRNA^{Leu} complex with paromomycin), 3UZ6 and 3UZ9 (cognate tRNA^{Tyr} complex), 3UZG and 3UZF (near-cognate tRNA^{Tyr} complex), and 3UZL and 3UZK (near-cognate tRNA^{Tyr} complex with paromomycin). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.Y. (marat@igbmc.fr) or G.Y. (gula@igbmc.fr).

METHODS

Complex formation and crystallization. Ribosomes were purified from *Thermus thermophilus* cells as described before²¹. The 30-nucleotide-long mRNA constructs I–IV (see below) were purchased from Dharmacon. In all sequences the AUG start codon is underlined and the A codons are in bold. Purified native uncharged *Escherichia coli* tRNA^{fMet}, tRNA₂^{Leu} and tRNA^{Tyr} were supplied by Chemical Block. Aminoglycoside antibiotic paromomycin was purchased from Sigma-Aldrich.

The cognate and near-cognate complexes were formed in 10 mM tris-acetate, pH 7.0, 40 mM KCl, 7.5 mM Mg(Ac)₂, 0.5 mM DTT by incubating 70S ribosomes (3 µM) with mRNA-I, -II, -III or -IV and tRNA^{fMet} for 10 min at 37 °C. Then tRNA₂^{Leu} and tRNA^{Tyr} were added to the mixtures with mRNA-I or -II and mRNA-III or -IV, respectively, and the complexes were further incubated for 30 min. For all complexes mRNA, tRNA^{fMet} and tRNA₂^{Leu} (or tRNA^{Tyr}) were present in fivefold stoichiometric excess of the ribosome concentration. Complexes with paromomycin were obtained by including the antibiotic (30 µM) into the incubation mixture containing mRNA-II with tRNA₂^{Leu} and mRNA-IV with tRNA^{Tyr}. Crystals were grown at 24 °C by sitting-drop vapour diffusion as described before²⁰. In agreement with previous results, initiator tRNA^{fMet} was found in the P site of all complexes and either tRNA₂^{Leu} or tRNA^{Tyr} was found in the A site (tRNA^{fMet} was easily distinguishable from tRNA₂^{Leu} and tRNA^{Tyr} based on the large variable loops in those tRNAs). In the complexes with mRNA III and IV, the E site was occupied by tRNA^{Tyr}. However, in complexes with mRNA I and II, the quality of the density did not allow identification of the E site tRNA which was then modelled as tRNA^{fMet}. mRNA-I 5'-GGCAAGGAGGU(U)₄AUGCUC(U)₉-3' (cognate for tRNA₂^{Leu}). mRNA-II 5'-GGCAAGGAGGU(U)₄AUGUUU(U)₉-3' (near-cognate for tRNA₂^{Leu}). mRNA-III 5'-GGCAAGGAGGU(A)₄AUGUAC(A)₉-3' (cognate for tRNA^{Tyr}). mRNA-IV 5'-GGCAAGGAGGU(A)₄AUGUGC(A)₉-3' (near-cognate for tRNA^{Tyr}).

Data collection, processing and structure determination. All crystals belong to space group P2₁2₁2₁ and contain two molecules per asymmetric unit. Data on all six complexes were collected at 100 K at the Synchrotron Light Source,

Switzerland, using the Pilatus 6M detector. A very low dose mode was used and huge redundancy was collected²⁹. The structure²¹, with tRNAs, mRNA and metal ions removed, was used for refinement with Phenix³⁰. The initial model was correctly placed within each data set by rigid body refinement with each molecule as a rigid body. This was followed by rigid body refinement of individual subunits. After positional and B-factor refinement, the resulting electron density maps were inspected and the tRNAs and mRNA ligands were built in these unbiased maps. In all of these and the following refinement rounds, no base-pair restraints were used between tRNAs and mRNAs to avoid bias towards perfect base-pair geometry. After several cycles of manual rebuilding followed by positional and individual isotropic B-factor refinement, magnesium ions were added and a final refinement round took place. A summary of the crystallographic data and refinement statistics is given in Supplementary Tables 1 and 2. Supplementary Table 3 shows the average B-factor for the entire structure as well as average B-factors for the decoding centre and the codon–anticodon helix. From this table it is seen that the average B-factors for the substructure comprising the decoding centre (G530, A1492, A1493 from 16S and G1913 from 23S), as well as the codon–anticodon helices (nucleotides 34–36 of tRNA in the A site and the corresponding codon of mRNA), are less than the overall B-factor for the entire ribosome structure. Therefore it is clear that the decoding centre is part of the most accurately determined parts of these models.

To verify that the base-pair geometry described in the paper is correct, we performed a many extra independent refinement rounds with base-pair geometries restrained to various standards (Watson–Crick, wobble, etc.) so that we could be confident about the reported geometries. Supplementary Figs 3 and 4 show OMIT-averaged kick maps^{30,31} of the G•U mismatch for the two near-cognate complexes. These unbiased maps show that a wobble conformation of these base pairs would not fit into the electron density and clearly demonstrate that a Watson–Crick conformation is the only plausible fit.

31. Praenikar, J., Afonine, P. V., Guncar, G., Adams, P. D. & Turk, D. Averaged kick maps: less noise, more signal and probably less bias. *Acta Crystallogr. D* **65**, 921–931 (2009).

Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes

Wen Tang^{1,2,3}, Ram Kannan^{1,2,3}, Marco Blanchette^{2,4} & Peter Baumann^{1,2,3}

In most eukaryotes, the progressive loss of chromosome-terminal DNA sequences is counteracted by the enzyme telomerase, a reverse transcriptase that uses part of an RNA subunit as template to synthesize telomeric repeats. Many cancer cells express high telomerase activity, and mutations in telomerase subunits are associated with degenerative syndromes including dyskeratosis congenita and aplastic anaemia. The therapeutic value of altering telomerase activity thus provides ample impetus to study the biogenesis and regulation of this enzyme in human cells and model systems. We have previously identified a precursor of the fission yeast telomerase RNA subunit (TER1)¹ and demonstrated that the mature 3'-end is generated by the spliceosome in a single cleavage reaction akin to the first step of splicing². Directly upstream and partly overlapping with the spliceosomal cleavage site is a putative binding site for Sm proteins. Sm and like-Sm (Lsm) proteins belong to an ancient family of RNA-binding proteins represented in all three domains of life³. Members of this family form ring complexes on specific sets of target RNAs and have critical roles in their biogenesis, function and turnover. Here we demonstrate that the canonical Sm ring and the Lsm2–8 complex sequentially associate with fission yeast TER1. The Sm ring binds to the TER1 precursor, stimulates spliceosomal cleavage and promotes the hypermethylation of the 5'-cap by Tgs1. Sm proteins are then replaced by the Lsm2–8 complex, which promotes the association with the catalytic subunit and protects the mature 3'-end of TER1 from exonucleolytic degradation. Our findings define the sequence of events that occur during telomerase biogenesis and characterize roles for Sm and Lsm complexes as well as for the methylase Tgs1.

In eukaryotes, seven Sm proteins (SmB, SmD1, 2 and 3, SmE, SmF and SmG) form a heteroheptameric complex at U-rich Sm-binding sites (AU_{4–6}GR) of various small nuclear RNAs (snRNAs) including the spliceosomal snRNAs U1, U2, U4 and U5 (refs 4, 5). Assembly of Sm proteins *in vivo* requires the help of the survival of motor neuron protein (SMN), mutations in which result in spinal muscular atrophy⁶. At least two Sm-like complexes have been characterized. The Lsm1–7 complex functions in messenger RNA (mRNA) degradation^{7,8} and the Lsm2–8 complex is involved in the maturation of various polymerase III transcripts^{9–11} and ribosomal RNAs¹². Purified Lsm2–8 binds to the 3'-terminal U-tract on U6, but not to the internal U-rich Sm sites in U1, U2, U4 and U5 snRNAs, illustrating that Sm and Lsm complexes have different sets of target RNAs⁹.

Sm-binding sites are also found near the 3'-ends of telomerase RNA subunits from diverse yeasts^{1,13–16} and are important for RNA processing and/or stability^{1,2,15}. Actual binding of Sm proteins has been demonstrated for TLC1, the telomerase RNA from *Saccharomyces cerevisiae*¹⁵, but the functional consequences of this interaction have remained largely unexplored. The Sm-binding site in TLC1 is located several nucleotides upstream of the mature 3'-end¹³. In contrast, spliceosomal cleavage of *Schizosaccharomyces pombe* TER1 truncates the putative Sm-binding site by one nucleotide², which may compromise stable association of the Sm ring with mature TER1. We therefore set

out to examine which proteins bind to the 3'-end of mature TER1, and to determine the function of the putative Sm site for TER1 biogenesis and stability.

A strategy was developed to examine the 3'-end of TER1 by massively parallel sequencing to obtain a quantitative measure of 3'-end sequence distribution and to identify the most abundant terminal sequences (Fig. 1a). This analysis revealed that, after spliceosomal cleavage, over 60% of TER1 molecules lost additional nucleotides at the 3'-end and terminate in a stretch of three to six uridines (Fig. 1b). The 3'-end of most of TER1 therefore resembles the 3'-end of U6 snRNA, which is bound by the Lsm2–8 complex. To test whether Sm or Lsm proteins associate with TER1, carboxy-terminal c-Myc epitope tags were inserted at the genomic loci of all Sm and Lsm proteins.

Immunoprecipitations were performed with a subset of strains that did not show overt growth defects, expressed tagged proteins and maintained telomeres (Supplementary Fig. 1). The snRNA U1 control specifically co-precipitated with Sm proteins, confirming that the epitope tags did not interfere with immunoprecipitation of RNP complexes (Fig. 1c). TER1 co-immunoprecipitated with all four Sm

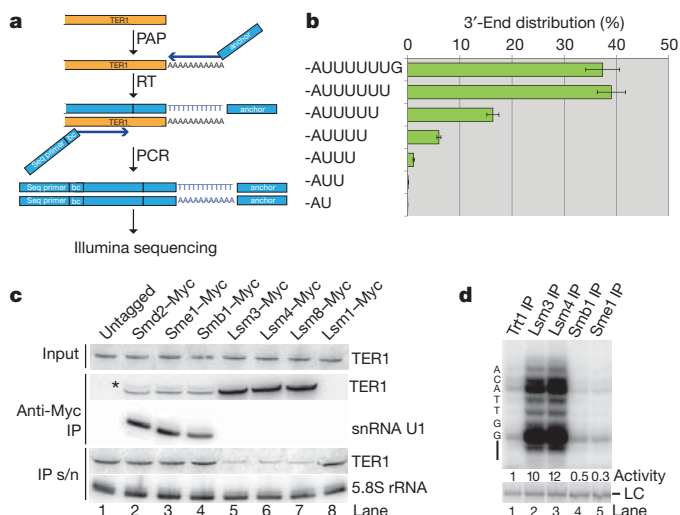


Figure 1 | TER1 RNA associates with Sm and Lsm proteins. **a**, Method used to map the 3'-end distribution of TER1 post spliceosomal cleavage. RNA is depicted in orange, DNA in blue. PAP, poly(A) polymerase; RT, reverse transcription; PCR, polymerase chain reaction; bc, barcode. **b**, Distribution of 3'-end positions in mature TER1 from wild-type cells. The average of four experiments is shown; error bars, standard deviation; a total of 23×10^6 sequences were scored. **c**, Northern blot of RNA isolated from immunoprecipitations with anti-c-Myc antibodies. Input and immunoprecipitation (IP) supernatant (s/n) represent 10% of the sample. An asterisk marks the position of the TER1 precursor. The lower band corresponds to the mature form of TER1. **d**, Telomerase activity assay performed on beads from immunoprecipitation of tagged proteins as indicated above each lane. Activity was quantified relative to the Trt1 immunoprecipitate sample. A 100-mer [³²P]oligonucleotide was used as recovery and loading control (LC).

¹Howard Hughes Medical Institute, Kansas City, Missouri 64110, USA. ²Stowers Institute for Medical Research, Kansas City, Missouri 64110, USA. ³Department of Molecular and Integrative Physiology, University of Kansas Medical Center, Kansas City, Kansas 66160, USA. ⁴Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, Kansas 66160, USA.

proteins tested (Fig. 1c, lanes 2–4, and Supplementary Fig. 2a), including members of each of the three Sm subcomplexes⁴. Strikingly though, several-fold more TER1 was recovered from Lsm immunoprecipitates resulting in an approximately 80% depletion of TER1 from the immunoprecipitation supernatant (Fig. 1c, lanes 5–7). TER1 precipitated with all subunits of the Lsm2–8 complex (Fig. 1c and Supplementary Fig. 2b), but not with Lsm1 (Fig. 1c, lane 8), the subunit specific to the Lsm1–7 complex.

To determine whether Sm and/or Lsm are associated with active telomerase, direct *in vitro* activity assays were performed on immunoprecipitates. Telomerase activity was detected in all samples, but was 20-fold higher in Lsm3 and 4 than Smb1 and Sme1 immunoprecipitates (Fig. 1d and Supplementary Fig. 2c). In part this can be explained by the lower recovery of telomerase with Sm proteins, as judged by quantification of telomerase RNA on northern blots (Supplementary Fig. 2c, d). However, even after normalization to the amount of TER1 in each immunoprecipitate, Lsm-associated telomerase activity was still 2.8-fold higher than that associated with Sm proteins. The simplest explanation for this observation is that a fraction of Sm-associated TER1 is not yet associated with the catalytic subunit of telomerase. Indeed, further experiments confirmed that Sm binding precedes Trt1 binding to TER1.

To gain insights into the functions of Sm and Lsm binding to telomerase we initially focused on the Sm association. For most characterized snRNAs, sequences downstream of the Sm-binding site are critical for Sm loading¹⁷. As the mature form of TER1 lacks such sequences, we tested whether the Sm complex was loaded onto the TER1 precursor before spliceosomal cleavage. Reverse transcription PCR (RT-PCR) confirmed that the precursor is indeed specifically associated with the Sm complex, but is undetectable in Lsm immunoprecipitations (Fig. 2a).

As the spliceosome contains Sm complexes, the TER1–Sm interaction may reflect binding of the spliceosome to the TER1 precursor. To test whether Sm proteins bind TER1 directly, we generated constructs with either a mutant 5′-splice-site or a deletion of the intron. Both mutant RNAs co-immunoprecipitated with Smb1 (Fig. 2b). In contrast, replacing the Sm-binding sequence with a random sequence (*ter1-sm6* mutant) reduced Sm association by 22-fold (Fig. 2c). Similarly, Lsm association was undetectable for *ter1-sm6* (Supplementary Fig. 3a). We therefore surmised that Sm and Lsm proteins directly bind to the previously identified site in TER1.

We next examined the effect of Sm binding on 3′-end processing by the spliceosome. Loss of Sm binding in the *ter1-sm6* mutant resulted in a sevenfold reduction in the processed form (Fig. 2d). Furthermore, a series of deletion mutants within the Sm site caused progressive inhibition of TER1 cleavage (Supplementary Fig. 3b), but not TER1 splicing (Supplementary Fig. 3c). Finally, introducing an eight-nucleotide spacer between the Sm site and 5′-cleavage-site also impaired processing (Fig. 2e). In summary, weakening or abolishing Sm association with the TER1 precursor reduces spliceosomal cleavage, indicating that Sm proteins promote 3′-end processing of TER1.

A conserved feature among yeast and mammalian telomerase RNAs is the post-transcriptional hypermethylation of the 5′-cap into a 2,2,7-trimethyl guanosine (TMG) form^{1,15,18}. Sm proteins were first implicated in promoting cap hypermethylation on U2 snRNA in *Xenopus* extract¹⁹. It was later shown *in vitro* that TMG-capping of human U1 requires the presence of Smb/B′-SmD3 (refs 4, 20). A screen for physical association with Sm proteins led to the identification of the methylase Tgs1 in budding yeast²¹. To elucidate the roles of Sm and/or Lsm in the hypermethylation of the 5′-cap on TER1, we tested which, if any, of these proteins interact with *S. pombe* Tgs1 (ref. 22) by two-hybrid analysis. Smd proteins scored positive, with Smd2 displaying the strongest interaction, and the other Sm proteins and all Lsm proteins showing weak or no interaction (Supplementary Fig. 4a). We next examined whether preventing Sm binding to TER1 affects cap hypermethylation. Whereas wild-type TER1 was readily precipitated

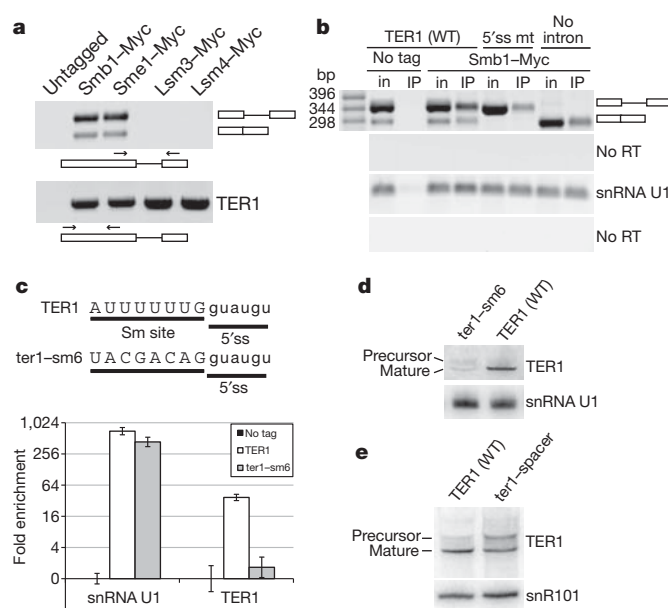


Figure 2 | Sm proteins associate with TER1 precursor and promote spliceosomal cleavage. **a**, RNA from anti-c-Myc immunoprecipitates was analysed by RT-PCR using primers in the first and second exon (primers represented by arrows in the schematic below the gel) to amplify the precursor form (upper panel). The primer pair also amplifies the spliced form (lower band in Sm immunoprecipitates). A primer pair in the first exon was used to visualize all forms of TER1 combined (lower panel). **b**, Sm association does not require spliceosome assembly on TER1. RT-PCR was performed on RNA purified from input (in) and anti-c-Myc immunoprecipitate beads (IP). Primers amplifying snRNA U1 were used as a positive control. **c**, The Sm-binding site (upper case) and 5′-splice-site (5′ss, lower case) for wild-type TER1 and the *ter1-sm6* mutant (MT). Replacing the Sm-binding site on TER1 (*ter1-sm6* mutant) compromises Sm association. RNA recovered from anti-c-Myc immunoprecipitates from untagged control and Smb1–Myc strains was quantified by real-time PCR. Data are plotted as enrichment over the untagged control. Error bars, standard error of triplicate experiments. **d**, Sm site mutation affects TER1 spliceosomal cleavage. Total RNA samples were analysed by northern blot for TER1 and snRNA U1. **e**, Increasing the distance between the Sm site and 5′-splice-site in the *ter1-spacer* mutant (AU₆GgccaauGUGU) impairs TER1 processing. Northern blot for TER1 and snRNA snR101 as loading control.

with a monoclonal antibody against the TMG cap, *ter1-sm6* recovery was at least 25-fold reduced (Fig. 3a and Supplementary Fig. 4b). Only the cleaved form of TER1 was recovered in TMG immunoprecipitations from wild-type cells, suggesting that spliceosomal cleavage precedes hypermethylation (Supplementary Fig. 4c). TER1 was not TMG-capped in a *tgs1Δ* strain, confirming that Tgs1 is the enzyme responsible for TER1 cap hypermethylation (Supplementary Fig. 4d).

In light of the reported increase in telomerase RNA in *tgs1Δ* budding yeast²³, we were surprised to observe a fivefold reduction in mature TER1 RNA in *tgs1Δ* compared with wild type in *S. pombe* (Fig. 3b). In addition, an increase in the precursor indicated a 3′-end processing defect. The viability of *tgs1Δ* cells ruled out a major splicing defect, but we consistently noted a small reduction in spliceosomal snRNAs isolated from *tgs1Δ* cells (Fig. 3b and data not shown). To differentiate between a processing defect and a direct effect of the TMG cap on TER1 stability, we mutated the spliceosomal cleavage site and inserted a hammerhead ribozyme sequence to generate the mutant *ter1-5′ssmut-HH* (Supplementary Fig. 4e). In this construct, processing of TER1 occurs independently of the spliceosome by ribozyme cleavage. When comparing *ter1-5′ssmut-HH* levels between wild-type and *tgs1Δ* cells, a twofold reduction was observed (Fig. 3b). Taken together, these results show that *tgs1Δ* affects TER1 processing by the spliceosome as well as TER1 stability. Consistent with the exquisite dosage sensitivity for telomerase RNA in diverse species^{24,25}, this reduction in TER1

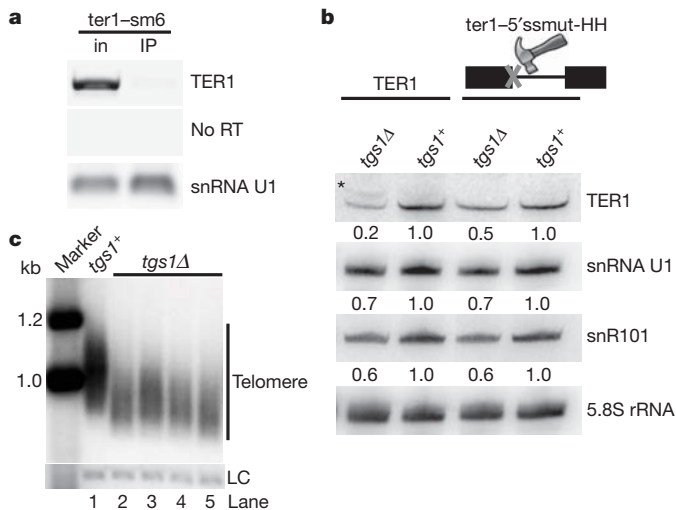


Figure 3 | Tgs1 modifies TER1 and is required for normal telomere maintenance. **a**, Loss of Sm site compromises TMG cap formation. RT-PCR amplifying all forms of TER1 and *ter1-sm6* mutant from anti-TMG immunoprecipitation (IP) and input (in) samples; snRNA U1 served as control. **b**, Bypass of spliceosomal cleavage reveals functions of Tgs1 in TER1 processing and stability. Northern blot analysis of TER1, snRNA U1, snR101 and 5.8S rRNA from total RNA prepared from wild-type and *tgs1Δ* strains harbouring either TER1 or the *ter1-5'ssmut-HH* mutant. An asterisk marks the position of the TER1 precursor. **c**, Deletion of *tgs1*⁺ causes telomere shortening. Telomere length was analysed by Southern blotting of EcoRI-digested genomic DNA from four independent *tgs1Δ* isolates and an otherwise isogenic *tgs1*⁺ strain. A probe for the *rad16* gene was used as a loading control (LC).

resulted in shorter telomeres (Fig. 3c). Neither telomerase activity nor Lsm association was reduced beyond the effects expected from the reduced steady-state level of TER1 (Supplementary Fig. 4f, g).

Most TER1 post-spliceosomal cleavage was bound by Lsm2–8, but a small fraction was associated with Sm proteins (Fig. 1c). To investigate whether this was indicative of a switch from Sm to Lsm binding, we examined the distribution of 3'-ends in each immunoprecipitation by massively parallel sequencing. Around 70% of Sm-bound TER1 post-cleavage terminated precisely at the spliceosomal cleavage site (Fig. 4a and Supplementary Fig. 5a). Enrichment of this form in the Sm-bound fraction is consistent with Sm proteins binding the TER1 precursor and remaining associated with TER1 until after cleavage and cap hypermethylation have occurred. In contrast, Lsm-associated TER1 predominantly terminated in U₃₋₆, indicating that a switch between Sm and Lsm binding occurs after spliceosomal cleavage and is associated with exonucleolytic processing (Fig. 4a and Supplementary Fig. 5b).

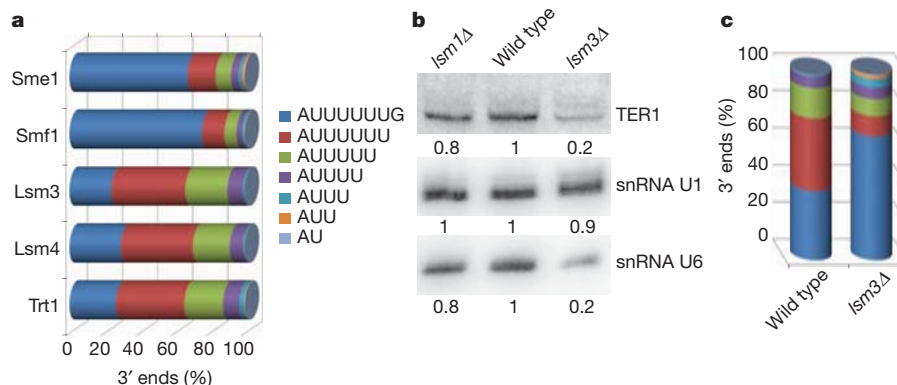


Figure 4 | Lsm proteins replace Sm and protect the 3'-end of TER1. **a**, 3'-End sequence distribution of TER1 from immunoprecipitation samples. **b**, Northern blot analysis from total RNA prepared from wild-type, *lsm1Δ* and *lsm3Δ* strains, quantified relative to wild type for each RNA. **c**, Specific loss of

Consistent with most telomerase activity being associated with Lsm2–8, the TER1 3'-end distribution from Trt1 immunoprecipitates was indistinguishable from that of Lsm-bound TER1.

The observation that loss of Sm binding coincided with the loss of terminal nucleotides led us to speculate that Lsm2–8 may function in protecting the 3'-end of TER1 against further exonucleolytic degradation. To test this, we attempted to generate Lsm deletion strains. Whereas most Lsm proteins are essential, *lsm1Δ* and *lsm3Δ* cells were viable. Consistent with a protective function for Lsm2–8, the levels of TER1 and U6 snRNA were reduced approximately fivefold in *lsm3Δ* cells (Fig. 4b). No such effect was seen when deleting *lsm1*, nor was the level of U1 snRNA reduced in *lsm3Δ* cells. The 3'-end sequence distribution for TER1 from total RNA of *lsm3Δ* cells closely resembled the Sm-bound fraction in wild type, whereas the Lsm-bound fraction was selectively lost in the mutant (Fig. 4c and Supplementary Fig. 5c). The viability of *lsm3Δ* cells further allowed us to confirm that cap hypermethylation is unaffected by the absence of Lsm consistent with Tgs1 acting on TER1 before Lsm binding (Supplementary Fig. 5d).

To verify independently a role for Lsm proteins in stabilizing TER1, we took advantage of the observation that Lsm binding requires a stretch of consecutive uridines⁹. In contrast, Sm binding tolerates other nucleotides in certain positions of the binding motif, as exemplified by the Sm-binding site in human U1 snRNA (AAUUUGUG). When the TER1 Sm site was mutated to reduce the number of consecutive uridines, the level of mature TER1 was decreased (Fig. 5a). We next precipitated Smb1, Lsm4 and Trt1 from wild type and strains containing the *ter1-SmU1* mutant. As expected, the mutation had little effect on the binding of Sm proteins (Fig. 5b). In fact, when normalized for the lower level of *ter1-SmU1* compared with wild type, recovery of *ter1-SmU1* with Smb1 was increased 1.6-fold. In contrast, Lsm binding was diminished by more than 20-fold. Most surprisingly, the interaction between the catalytic subunit Trt1 and telomerase RNA was also compromised in the *ter1-SmU1* mutant (Fig. 5b). The normalized recovery of *ter1-SmU1* with Trt1 was 15-fold lower than wild type, indicating that Lsm binding facilitates Trt1–TER1 association, possibly by inducing a conformational change in the RNA analogous to how binding of the p65 protein facilitates telomerase assembly in *Tetrahymena*^{26–28}. Consistent with the poor recovery of *ter1-SmU1* in Trt1 immunoprecipitations, *in vitro* telomerase activity was below the level of detection (Fig. 5c).

Analysis of the 3'-end sequence distribution for *ter1-SmU1* from total RNA revealed that most of the mutant RNA ends at the cleavage site (Supplementary Fig. 6). This form constituted close to 90% of *ter1-SmU1* in Smb1 immunoprecipitates. In contrast, Lsm4 and Trt1 immunoprecipitates predominantly recovered RNA ending in -AUUU and -AUUUG (Supplementary Fig. 6). These results further support that Trt1 preferentially associates with Lsm-bound telomerase RNA.

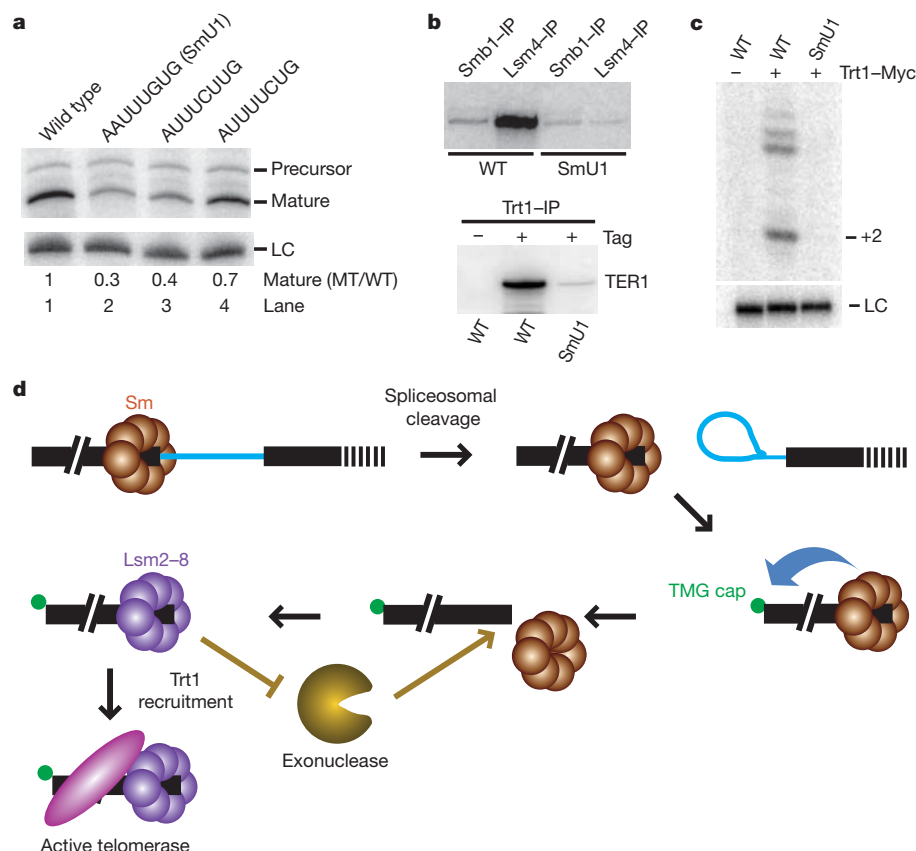


Figure 5 | Lsm binding to TER1 promotes telomerase assembly and protects TER1 from degradation. **a**, Northern blot for TER1. The indicated ratios of mutant (MT) to wild type (WT) are normalized to the loading control snR101 (LC). **b**, Northern blot for TER1 and the ter1-SmU1 mutant using RNA isolated from anti-c-Myc immunoprecipitations performed on extract from

strains harbouring Smb1-Myc, Lsm4-Myc or Trt1-Myc as indicated. **c**, Telomerase activity assay performed on Trt1 immunoprecipitates from strains harbouring either wild type or ter1-SmU1. An untagged Trt1 strain was used as negative control. **d**, Sequence of events that occur during telomerase biogenesis.

They also confirm the role of Lsm in protecting the 3'-end of TER1 from further degradation, as diminished Lsm binding coincides with an overall reduction in telomerase RNA and a shift towards the form that is bound by Sm.

Taken together, our observations demonstrate that distinct populations of TER1 molecules associate with the Sm and Lsm complexes and suggest a sequence of events for TER1 biogenesis (Fig. 5d). The polyadenylated TER1 precursor is bound by the Sm complex, which promotes spliceosomal cleavage and subsequent 5'-cap hypermethylation by recruiting Tgs1. The Sm ring is then replaced by the Lsm2-8 complex, which protects TER1 from exonucleolytic degradation and promotes binding of the catalytic subunit.

Despite their structural similarity and related binding motifs, Sm and Lsm complexes have different modes of RNA binding and were thought to have distinct and non-overlapping sets of target RNAs. The finding that the TER1 precursor is exclusively associated with the Sm complex, whereas most mature TER1 is bound by Lsm2-8, revealed that biogenesis of telomerase RNA involves both Sm and Lsm complexes. Considering the central roles that Sm and Lsm proteins play in RNA metabolism, it will be important to determine whether biogenesis of other non-coding RNAs also involves Sm- and Lsm2-8-bound stages. Furthermore, it is interesting to note that several human Sm/Lsm proteins have been reported to co-purify with telomerase^{29,30}, raising the possibility that these proteins also function in TMG cap formation and telomerase assembly in metazoans.

METHODS SUMMARY

Myc epitope tags were integrated at the genomic loci and immunoprecipitations were performed in whole-cell extracts with anti-c-Myc antibodies. The different forms of telomerase RNA were detected by northern blotting and RT-PCR. The

distribution of 3'-ends was assessed at single nucleotide resolution by preparing libraries of oligo(A)-tailed telomerase RNA and massively parallel sequencing.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 September 2011; accepted 7 February 2012.

Published online 25 March 2012.

- Leonardi, J., Box, J. A., Bunch, J. T. & Baumann, P. TER1, the RNA subunit of fission yeast telomerase. *Nature Struct. Mol. Biol.* **15**, 26–33 (2008).
- Box, J. A., Bunch, J. T., Tang, W. & Baumann, P. Spliceosomal cleavage generates the 3'-end of telomerase RNA. *Nature* **456**, 910–914 (2008).
- Wilusz, C. J. & Wilusz, J. Eukaryotic Lsm proteins: lessons from bacteria. *Nature Struct. Mol. Biol.* **12**, 1031–1036 (2005).
- Raker, V. A., Plessel, G. & Lührmann, R. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. *EMBO J.* **15**, 2256–2269 (1996).
- Patel, S. B. & Bellini, M. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* **36**, 6482–6493 (2008).
- Terns, M. P. & Terns, R. M. Macromolecular complexes: SMN — the master assembler. *Curr. Biol.* **11**, R862–R864 (2001).
- Tharun, S. *et al.* Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**, 515–518 (2000).
- Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M. & Seraphin, B. A Sm-like protein complex that participates in mRNA degradation. *EMBO J.* **19**, 1661–1671 (2000).
- Achsel, T. *et al.* A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation *in vitro*. *EMBO J.* **18**, 5789–5802 (1999).
- Mayes, A. E., Verdone, L., Legrain, P. & Beggs, J. D. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J.* **18**, 4321–4331 (1999).
- Kufel, J., Allmang, C., Verdone, L., Beggs, J. D. & Tollervey, D. Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. *Mol. Cell. Biol.* **22**, 5248–5256 (2002).
- Kufel, J., Allmang, C., Petfalski, E., Beggs, J. & Tollervey, D. Lsm proteins are required for normal processing and stability of ribosomal RNAs. *J. Biol. Chem.* **278**, 2147–2156 (2003).

13. Dandjinou, A. T. *et al.* A phylogenetically based secondary structure for the yeast telomerase RNA. *Curr. Biol.* **14**, 1148–1158 (2004).
14. Gunisova, S. *et al.* Identification and comparative analysis of telomerase RNAs from *Candida* species reveal conservation of functional elements. *RNA* **15**, 546–559 (2009).
15. Seto, A. G., Zaug, A. J., Sobel, S. G., Wolin, S. L. & Cech, T. R. *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature* **401**, 177–180 (1999).
16. Webb, C. J. & Zakian, V. A. Identification and characterization of the *Schizosaccharomyces pombe* TER1 telomerase RNA. *Nature Struct. Mol. Biol.* **15**, 34–42 (2008).
17. Yong, J., Kasim, M., Bachorik, J. L., Wan, L. & Dreyfuss, G. Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol. Cell* **38**, 551–562 (2010).
18. Jady, B. E., Bertrand, E. & Kiss, T. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J. Cell Biol.* **164**, 647–652 (2004).
19. Mattaj, J. W. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell* **46**, 905–911 (1986).
20. Plessel, G., Fischer, U. & Lührmann, R. m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) *in vitro*: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. *Mol. Cell Biol.* **14**, 4160–4172 (1994).
21. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. & Bordonne, R. Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell* **9**, 891–901 (2002).
22. Hausmann, S., Ramirez, A., Schneider, S., Schwer, B. & Shuman, S. Biochemical and genetic analysis of RNA cap guanine-N2 methyltransferases from *Giardia lamblia* and *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **35**, 1411–1420 (2007).
23. Franke, J., Gehlen, J. & Ehrenhofer-Murray, A. E. Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. *J. Cell Sci.* **121**, 3553–3560 (2008).
24. Harrington, L. Making the most of a little: dosage effects in eukaryotic telomere length maintenance. *Chromosome Res.* **13**, 493–504 (2005).
25. Mozdy, A. D. & Cech, T. R. Low abundance of telomerase in yeast: implications for telomerase haploinsufficiency. *RNA* **12**, 1721–1737 (2006).
26. Berman, A. J., Gooding, A. R. & Cech, T. R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol. Cell Biol.* **30**, 4965–4976 (2010).
27. O'Connor, C. M. & Collins, K. A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. *Mol. Cell Biol.* **26**, 2029–2036 (2006).
28. Stone, M. D. *et al.* Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458–461 (2007).
29. Fu, D. & Collins, K. Human telomerase and Cajal body ribonucleoproteins share a unique specificity of Sm protein association. *Genes Dev.* **20**, 531–536 (2006).
30. Fu, D. & Collins, K. Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Mol. Cell* **28**, 773–785 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Shuman for the *tgs1Δ* strain, J. A. Box, J. T. Bunch, S. Hartnett and R. M. Helston for technical assistance, the Molecular Biology Core Facility for site-directed mutagenesis and sequencing, D. P. Baumann and R. M. Helston for proofreading the manuscript, and all members of the Baumann laboratory for discussions. This work was funded in part by the Stowers Institute for Medical Research. R.K. is supported by an award from the American Heart Association, and P.B. is an Early Career Scientist with the Howard Hughes Medical Institute.

Author Contributions P.B. and W.T. conceived the study and designed the experiments; W.T. performed most of the experiments with some assistance from those acknowledged and P.B.; R.K. contributed to the characterization of Sm mutants and analysed telomere length of Myc-tagged strains. M.B. wrote the script for sequence data analysis and provided advice; W.T., R.K. and P.B. analysed the data, and W.T. and P.B. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.B. (peb@stowers.org).

METHODS

Yeast strains and constructs. The genotypes of all strains used in this study are listed in Supplementary Table 1. Strains expressing c-Myc-tagged Sm and Lsm proteins were constructed in strain PP138 as described³¹. Mutants *ter1-sm6*, *ter1-smΔG*, *ter1-smΔUG*, *ter1-smΔU₂G*, *ter1-smΔU₃G* and *ter1-5' smut-HH* were integrated at the *ter1* genomic locus by gene replacement. Other *ter1* mutants were generated in the context of plasmid pJW10 using the QuikChange II XL site-directed mutagenesis kit (Stratagene) and introduced into PP407, PP694 or PP695 as described³¹.

Yeast two-hybrid analysis. Yeast two-hybrid analysis used the Matchmaker GAL4 Two Hybrid System 3 (Clontech). Briefly, *fgs1*⁺ cDNA was cloned into the vector pGBKT7, and each full-length *lsm* and *sm* cDNA was cloned into pGADT7. Plasmids were co-transformed into the yeast strain AH109 and positive transformants were selected on SD–Leu–Trp plates. Interactions were analysed by plating threefold serial dilutions of overnight cultures onto SD–Leu–Trp–His–Ade plates. Plates were incubated for three days at 30 °C.

Telomere length analysis and telomerase activity assay. Cells were propagated for at least 80–100 generations and telomere length was analysed by Southern blotting as described³². Telomerase activity assays were performed on Sepharose beads as described³³ after immunoprecipitation from cell extracts of strains harbouring Myc-tagged Trt1, Sm or Lsm proteins.

Immunoprecipitation and RNA isolation. *S. pombe* cells were grown in yeast extract supplements³² and 6 l of cell suspension were collected by centrifugation at a density of 5×10^6 cells per millilitre. Cells were washed in TMG(300) (10 mM Tris–HCl, pH 8.0, 1 mM magnesium chloride, 10% (v/v) glycerol, 300 mM sodium acetate), the pellet was resuspended in two packed cell volumes of TMG(300) plus supplements ($5 \mu\text{g ml}^{-1}$ chymostatin, $5 \mu\text{g ml}^{-1}$ leupeptin, $1 \mu\text{g ml}^{-1}$ pepstatin, 1 mM benzamidine, 1 mM DTT, 1 mM EDTA and 0.5 mM PMSF) and the suspension was frozen in liquid nitrogen. Cells were lysed under liquid nitrogen in a 6850 cryogenic mill (SPEX CertiPrep) with eight 2 min cycles at an impactor rate of 10 per second and a 2 min cooling time between cycles. The lysed cell powder was transferred into a 50 ml tube and allowed to thaw on ice for 30 min. Cell extracts were cleared by two rounds of centrifugation at 14,000g for 7 min and frozen in liquid nitrogen for storage at -80°C . The concentration of proteins in the whole-cell extract was measured by Bradford protein assay. For c-Myc immunoprecipitation, monoclonal anti-c-Myc antibody (20 μg , Sigma) was incubated with 150 μl protein A/G agarose slurry (Calbiochem) in phosphate buffered saline at room temperature for 30 min. Beads were washed three times with TMG(300) plus supplements and whole-cell extract (1.2 ml) was added at a concentration of 5 mg ml^{-1} together with RNasin (40 U, Promega), Tween 20 (0.1%) and heparin (1 mg ml^{-1}). For immunoprecipitation of TMG-capped RNAs, anti-TMG antibody (3 μg , Calbiochem) was bound to 50 μl protein A/G agarose slurry (Calbiochem), washed with TMG(300) and 150 μg total *S. pombe* RNA was added in 0.7 ml TMG(300). Samples were incubated on a rotator at 4°C for 4 h, then washed three times with TMG(300) plus supplements and 0.1% Tween 20 and once with TMG(50) (as TMG(300) but only 50 mM sodium acetate). Protease inhibitors were omitted for TMG immunoprecipitations. RNA was isolated by treatment with proteinase K (2.0 mg ml^{-1} in 0.5% (w/v) SDS, 40 mM EDTA, 20 mM Tris–HCl, pH 7.5) at 50°C for 15 min, followed by extraction with acidic phenol and ethanol precipitation. RNA was then analysed by northern blotting, RT–PCR and 3'-end sequencing.

RNA analysis. RNA isolation and northern blotting were performed as described² except that Biotyne Nylon Transfer Membrane (Pall Corporation) was used and samples shown in Fig. 5a were treated with RNaseH in the presence of oligonucleotides BLoli1043 (AGGCAGAAGACTCACGTACACTGCAC), BLoli1275 and PBoli560 (GCGGAATTCT₁₈) to obtain better separation of precursor and mature form. The TER1 probe was generated as described²; other RNAs were detected using 5'-[³²P]DNA oligonucleotides as follows: GCTGCAGAACTCATGCCAGGTA AGT (snRNA U1), CGCTATTGTATGGGGCCTTTAGATTCTTA (snRNA snR101), CTTTCATCGATGCGAGAGCAAGAGATCCGT (5.8S rRNA) and GCAGTGTATCTTGTGCAGGGGCCA (snRNA U6).

Semi-quantitative RT–PCR was performed as described previously² with primers BLoli1275 (CGGAAACGGAATTCAGCATGT) and BLoli1020 (CAAA CAATAATGAACGTCCTG) amplifying the intron-spanning region, and PBoli918 (ACAACGGACGAGCTACACTC) and BLoli1006 (CATTTAAGTGC TTGTCAGATCACAACG) amplifying a region in the first exon. BLoli2051 (GACCTTAGCCAGTCCACAGTTA) and BLoli2101 (ACCTGGCATGAGTTTC TGC) were used to amplify snRNA U1.

For quantitative real-time RT–PCR, reverse transcription for input and immunoprecipitated RNA were performed with antisense primer BLoli2860 (TGCTCAGACCAAGTGAAAAA) and BLoli2051. Real-time PCR was performed in triplicate 12.5 μl reactions using Power SYBR Green PCR Master Mix (Applied Biosystems) according to the manufacturer's instructions. BLoli2860 and BLoli2859 (GGATCAAAGCTTTTGCTTGT) were used to amplify the first exon of TER1. BLoli2051 and BLoli2101 were used to amplify snRNA U1. The qRT–PCR results were imported into Microsoft Excel and the average value and standard deviation of triplicate cycle threshold (C_t) values were calculated. Enrichment of immunoprecipitation is represented by ΔC_t (C_t value (immunoprecipitation sample) minus C_t value (input)) relative to the untagged control samples. Error bars in the graph represent the positive and negative range of the standard error of the mean.

3'-end cloning. DNase-treated total RNA samples (2.5 μg) or immunoprecipitated and purified RNA was incubated with poly(A) polymerase (600 U, US Biologicals), RNase inhibitor (RNasin, 40 U) and ATP (0.5 mM) in 20 μl reactions at 30°C for 30 min. The reaction volume was increased to 35.5 μl by the addition of the oligonucleotide BLoli2327 (CAAGCAGAAGACGGCATACGA(T)₁₈, 125 pmol) and dNTP mix (25 nmol), and reactions were incubated at 65°C for 3 min followed by slow cooling to room temperature. The reaction volume was then adjusted to 50 μl with first strand buffer (Invitrogen), dithiothreitol (5 mM), RNasin (40 U) and Superscript III reverse transcriptase (200 U, Invitrogen), and reactions were incubated at 50°C for 60 min. RNaseH (5 U, NEB) was added and incubation was continued at 37°C for 20 min. Aliquots (3 μl) of this reaction were used in PCR with Taq polymerase (5 U, NEB), primers (GTTTCAGAGTTCTACAGTCCGAC GATC##GCAAAATGTTAAAAGGAACG) and BLoli2330 (CAAGCAGAAGAC GGCATACGA) (200 nM each, ## represents a two-nucleotide barcode used for multiplexing) under the following conditions: 3 min at 94°C followed by 10 cycles of 30 s at 94°C , 45 s at 55°C and 60 s at 72°C , followed by 7 min at 72°C . PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and eluted with 46 μl elution buffer. In the second round of PCR, 23 μl of the eluted product was amplified with BLoli2329 (AATGATACGGCGACCAACCGACAGG TTCAGAGTTCTACAGTCCGA) and BLoli2330 (200 nM each) under the following conditions: 3 min at 94°C followed by 29 cycles of 30 s at 94°C , 45 s at 55°C and 60 s at 72°C , followed by 7 min at 72°C . PCR products were separated by electrophoresis on 1.5% agarose gels, and bands of the correct size were excised and purified. The concentration of the PCR products was measured using an Agilent 2100 Bioanalyzer (Agilent Technologies) and further adjusted to 10 nM for massively parallel sequencing using Illumina sequencing technology. Reads were analysed using a custom script written in BioPerl to filter for those that contained the TER1 sequence (GCAAAAN₁₀AACG) and to sort the reads into different bins based on the two-nucleotide barcodes. The nucleotide sequence between GCAAAAN₁₀AACG and the oligo(A) sequence resulting from the poly(A) polymerase treatment represents the end of TER1 and was used to determine the 3'-end sequence distribution at single nucleotide resolution. Further analysis and graphs were prepared in Microsoft Excel.

31. Bahler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**, 943–951 (1998).
32. Bunch, J. T., Bae, N. S., Leonardi, J. & Baumann, P. Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability. *Mol. Cell. Biol.* **25**, 5567–5578 (2005).
33. Haering, C. H., Nakamura, T. M., Baumann, P. & Cech, T. R. Analysis of telomerase catalytic subunit mutants in vivo and in vitro in *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. USA* **97**, 6367–6372 (2000).

Visualizing molecular juggling within a B₁₂-dependent methyltransferase complex

Yan Kung¹, Nozomi Ando^{1,2}, Tzanko I. Doukov^{1†}, Leah C. Blasiak^{1†}, Güneş Bender³, Javier Seravalli⁴, Stephen W. Ragsdale³ & Catherine L. Drennan^{1,2,5}

Derivatives of vitamin B₁₂ are used in methyl group transfer in biological processes as diverse as methionine synthesis in humans and CO₂ fixation in acetogenic bacteria^{1–3}. This seemingly straightforward reaction requires large, multimodular enzyme complexes that adopt multiple conformations to alternately activate, protect and perform catalysis on the reactive B₁₂ cofactor. Crystal structures determined thus far have provided structural information for only fragments of these complexes^{4–12}, inspiring speculation about the overall protein assembly and conformational movements inherent to activity. Here we present X-ray crystal structures of a complete 220 kDa complex that contains all enzymes responsible for B₁₂-dependent methyl transfer, namely the corrinoid iron-sulphur protein and its methyltransferase from the model acetogen *Moorella thermoacetica*. These structures provide the first three-dimensional depiction of all protein modules required for the activation, protection and catalytic steps of B₁₂-dependent methyl transfer. In addition, the structures capture B₁₂ at multiple locations between its ‘resting’ and catalytic positions, allowing visualization of the dramatic protein rearrangements that enable methyl transfer and identification of the trajectory for B₁₂ movement within the large enzyme scaffold. The structures are also presented alongside *in crystallo* spectroscopic data, which confirm enzymatic activity within crystals and demonstrate the largest known conformational movements of proteins in a crystalline state. Taken together, this work provides a model for the molecular juggling that accompanies turnover and helps explain why such an elaborate protein framework is required for such a simple, yet biologically essential reaction.

B₁₂-dependent methyl transfer lies at the heart of methylation biochemistry and is an essential reaction in human health and microbial CO₂ sequestration^{2,3}. In humans, methionine synthase (MetH) methylates homocysteine to form methionine to maintain cellular pools of folate (vitamin B₉) and S-adenosylmethionine (AdoMet), the universal methyl donor. MetH mutation or vitamin B₁₂ deficiency can cause serious health consequences, including megaloblastic anaemia and birth abnormalities such as neural tube defects¹³. Acetogenic bacteria, including *M. thermoacetica*, use the corrinoid iron-sulphur protein (CFeSP) and its methyltransferase (MeTr) together to catalyse methyl transfer in the Wood–Ljungdahl carbon fixation pathway for growth on CO₂ as the sole carbon source¹⁴.

For both MetH and CFeSP/MeTr, methyltetrahydrofolate (CH₃-H₄folate) is the methyl donor, and a protein-bound B₁₂ derivative (cobalamin for MetH and 5'-methoxybenzimidazolyl cobamide for CFeSP) is the methyl carrier. In acetogenic bacteria, the CH₃-H₄folate methyl group is derived from enzymatic reduction of CO₂, whereas in humans, CH₃-H₄folate is the predominant circulating form of the vitamin. Although CH₃-H₄folate is the common methyl source, methyl removal from the N⁵ tertiary amine is chemically challenging because the product, tetrahydrofolate (H₄folate), is a poor leaving

group¹. Therefore, a particularly powerful nucleophile is required, and B₁₂ with cobalt in the +1 oxidation state, a Co(I) species dubbed a ‘supernucleophile’¹⁵, is recruited. Such strong reactivity comes at a price: reducing the inactive Co(II) state to active Co(I) is thermodynamically challenging, as the Co(II/I) reduction potential is one of the lowest in nature, –504 mV in CFeSP and –526 mV in MetH^{16,17}. In CFeSP, an electron is first delivered from a partner protein to an Fe₄S₄ cluster harboured by an activation domain^{18,19}. The electron is then passed to Co(II) to yield Co(I) (equation (1)), which attacks CH₃-H₄folate to form CH₃-Co(III) (equation (2)). CFeSP then delivers the methyl group to the Ni₂Fe₄S₄ active site metallocluster (A-cluster) of acetyl-CoA synthase (ACS), where it becomes the methyl of acetyl-CoA, and B₁₂ returns to its nucleophilic Co(I) state.



During the catalytic cycle of both MetH and CFeSP/MeTr (Supplementary Fig. 1), a series of ‘molecular juggling’ acts must be performed in which domains rearrange to contact the B₁₂ cofactor. Crystal structures of a MetH B₁₂-binding fragment⁴ and CFeSP from *Carboxydothermus hydrogenoformans* (ChCFeSP)⁹ both depict a ‘resting’ state, where B₁₂ is buried by a protective ‘capping’ domain, shielded from unwanted chemistry but inaccessible to substrate. Because methyl transfer uses S_N2 substitution²⁰, large conformational changes must ‘uncap’ B₁₂ before chemistry can occur. B₁₂ is ‘uncapped’ in structures of MetH fragments that depict B₁₂ activation^{7,11,12}, but no structure has been solved that shows B₁₂- and CH₃-H₄folate-binding domains together to illustrate methyl transfer.

To visualize this elusive methyl transfer complex, we determined a 2.38 Å resolution structure of folate-free CFeSP/MeTr from *M. thermoacetica* (Fig. 1 and Supplementary Table 1). The homodimeric MeTr component (58 kDa) is virtually identical to previous structures of both MeTr^{6,10} (Supplementary Fig. 2a), root mean squared deviation (r.m.s.d.) for Cα atoms 0.39 Å, and the analogous MetH domain that binds CH₃-H₄folate⁸, r.m.s.d. 1.03–1.08 Å. MeTr and MetH both use (β/α)₈ triosephosphate isomerase (TIM) barrels to bind and activate CH₃-H₄folate for nucleophilic attack. Two CFeSPs are present in the complex, each containing two subunits. The small subunit (35 kDa) is a TIM barrel which acts as the B₁₂ ‘cap’ in the ChCFeSP structure⁹, while the large subunit (48 kDa) has three domains joined by linkers: an amino (N)-terminal Fe₄S₄ activation domain (residues 1–57), a TIM barrel domain (residues 93–312) and a carboxy (C)-terminal B₁₂-binding domain (residues 325–446). With the exception of the Fe₄S₄ and B₁₂ domains, discussed below, both CFeSP copies align well to the ChCFeSP structure, r.m.s.d. 0.81–0.85 Å (Supplementary Fig. 2b).

¹Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ²Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA. ⁴Department of Biochemistry, University of Nebraska, Lincoln, Nebraska 68588, USA. ⁵Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. †Present addresses: Stanford Synchrotron Radiation Lightsource, Menlo Park, California 94025, USA (T.I.D.); Institute of Marine and Environmental Technology, University of Maryland Center for Environmental Science, Baltimore, Maryland 21202, USA (L.C.B.).

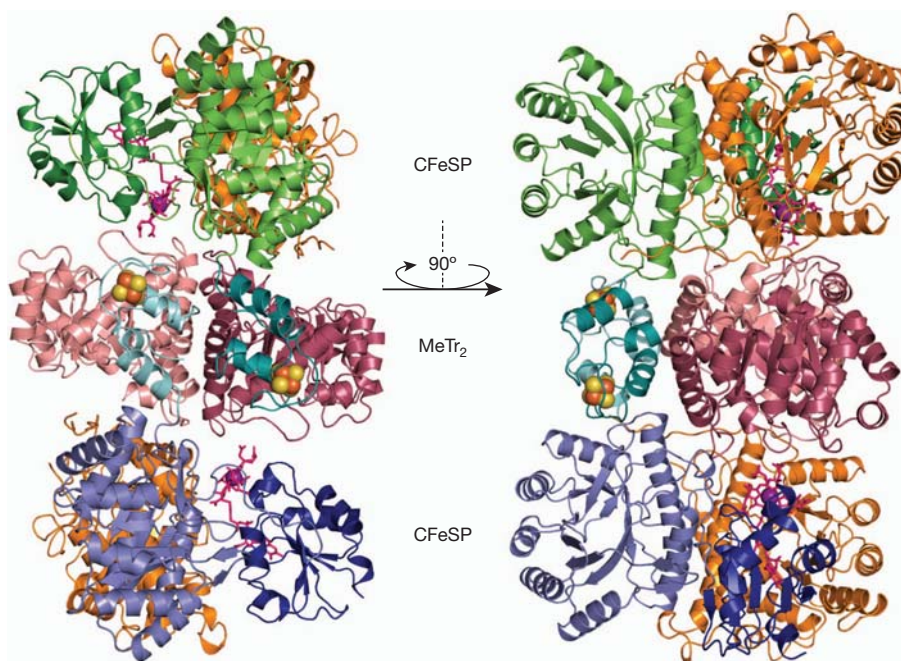


Figure 1 | The overall CFeSP/MeTr complex. Ribbon representation of MeTr homodimer (MeTr₂) in light and dark pink, CFeSP small subunits in orange, CFeSP large subunit Fe₄S₄ domains in teal and cyan, TIM barrel domains in

green and blue, and B₁₂ domains in dark green and dark blue. B₁₂ cofactors in magenta sticks with cobalt as violet spheres. Fe₄S₄ clusters in spheres: Fe in orange, S in yellow.

In the 220 kDa CFeSP/MeTr assembly (Fig. 1), the MeTr homodimer lies in the centre, with one CFeSP bound on either side. Each MeTr monomer has a C-terminal α -helix (residues 255–262) protruding from the TIM barrel rim. Contacts between this helix and its preceding loop with a CFeSP small subunit helix (residues 191–204) form the primary interactions between MeTr and CFeSP (Supplementary Fig. 3). Weak interactions between MeTr and CFeSP Fe₄S₄ domains have stabilized these highly flexible⁹ domains responsible for B₁₂ activation, allowing their visualization as bundles of short α -helices connected by long loops that coordinate the Fe₄S₄ cubane (Fig. 1 and Supplementary Figs 4 and 5). The Fe₄S₄ domains are observed to adopt a variety of positions that are all too far from the B₁₂ to afford reductive activation⁹. However, the long and primarily unstructured protein linkers that connect both the Fe₄S₄ and B₁₂ domains to the central TIM barrel must allow for the requisite flexibility for B₁₂ activation (Supplementary Figs 2b and 6).

B₁₂ domains of both MetH and CFeSP adopt Rossmann-like architectures that bind B₁₂ in the base-off conformation (Fig. 1)^{4,9}. High *B*-factors support the notion of flexibility mentioned above (Supplementary Fig. 7 and Supplementary Table 2), where electron density for both B₁₂ domains represents a highest occupancy position within an ensemble, rather than a sole conformation. In both CFeSPs, the average B₁₂ domain position resides between the ‘capping’ small subunit TIM barrel and the TIM barrel of a MeTr monomer, which are adjacent and nearly perpendicular to each other (Fig. 1). On average, the B₁₂ Co has shifted approximately 6.5 Å away from its ‘resting’ location towards the MeTr folate-binding site. B₁₂ in this structure is thus positioned ‘en route’ towards catalysis, with approximately 18 Å remaining to the methyl group of folate modelled into the MeTr active site, based on an alignment with the folate-bound MeTr structure¹⁰. In transitioning between ‘resting’ and ‘en route’ positions, the B₁₂ corrin ring breaks three interactions with the ‘capping’ domain and forms new contacts, including an H-bond with Asn 203 of MeTr (Fig. 2a, b).

Given the flexibility suggested by this structural analysis, we explored whether the B₁₂ domain can sample the 18 Å necessary to afford turnover within intact crystals, using anaerobic *in crystallo* ultraviolet–visible absorption spectroscopy to monitor the state of B₁₂. *In crystallo* and analogous solution spectra were collected in

parallel (Fig. 3) for the as-isolated Co(II) form of B₁₂. Reduction to Co(I) and methylation to CH₃-Co(III) were then achieved *in crystallo* and in solution, with all spectra matching well-established CFeSP absorption features^{18,19,21–23}. Importantly, these features disappear when light is passed through the solution surrounding the crystals, indicating that spectra represent protein in crystals and not protein that may have been liberated into the solution. Collectively, these data demonstrate enzymatic transfer of the CH₃-H₄folate methyl group to CFeSP-bound B₁₂, evidence that the B₁₂ domain is able to move at least 18 Å to trigger methyl transfer within the crystal. To our knowledge, this conformational movement represents the largest observed in a crystallized protein (Supplementary Discussion). Such dramatic B₁₂ domain movement is probably facilitated by the fact that CFeSP/MeTr is mostly composed of rigid TIM barrels that provide all the lattice contacts (Supplementary Fig. 8). Although their biosynthesis is energetically expensive, these high molecular mass TIM barrel scaffolds may be important for B₁₂-dependent methyltransferases to maintain structural integrity during the conformational gymnastics that alternately enable activation, protection and catalysis of the highly reactive B₁₂ cofactor. Thus, despite the small size of the transferred methyl moiety, these large conformational changes appear to necessitate large enzyme sizes.

Although the folate-free CFeSP/MeTr structure describes large B₁₂ domain movements that ‘uncap’ B₁₂ from the small subunit, it is interesting to consider why binding of CFeSP to MeTr does not simply position the B₁₂ domain directly over the MeTr active site. One explanation is that the structure represents an inactive complex; however, *in crystallo* results clearly demonstrate that CFeSP/MeTr crystals are active. Another explanation posits that an ensemble of ‘en route’ conformations exists when CH₃-H₄folate is absent, and that CH₃-H₄folate binding would shift the conformational equilibrium, moving B₁₂ closer to the folate-binding site. To obtain experimental support for this hypothesis, we solved additional CFeSP/MeTr structures co-crystallized with CH₃-H₄folate, with and without Ti(III) citrate as a reductant at 3.03 Å and 3.50 Å resolution, respectively. Absorption spectroscopy performed on these crystals shows that these structures represent a substrate form (CH₃-H₄folate bound, B₁₂ in the Co(II) state) and a product form (H₄folate bound, B₁₂ in the CH₃-Co(III) state) of the complex (Supplementary Fig. 9). Compared with

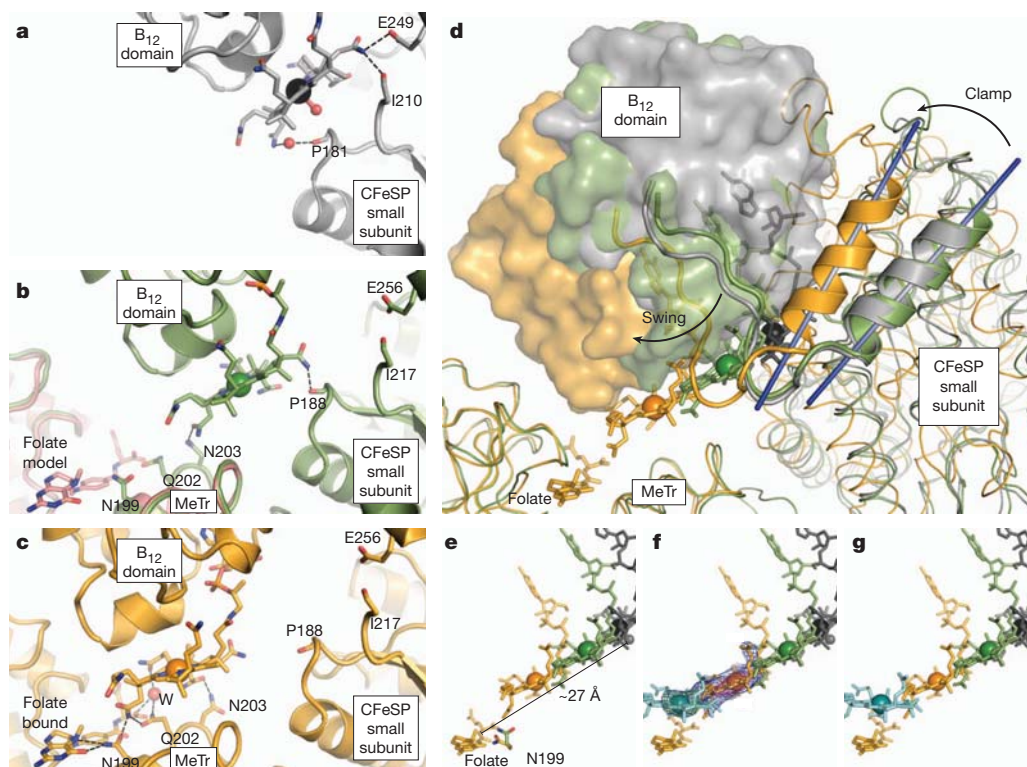


Figure 2 | Comparison of B₁₂ positions in 'resting' ChCFeSP, folate-free and folate-bound CFESP/MeTr. **a**, ChCFeSP (grey ribbons, Co of B₁₂: black sphere). **b**, Folate-free CFESP/MeTr (green ribbons, Co of B₁₂: green sphere) superimposed with CH₃-H₄folate-bound MeTr (Protein Data Bank accession number 2E7F, pink ribbons). **c**, Folate-bound CFESP/MeTr (orange ribbons, Co of B₁₂: orange sphere). Parts **a**–**c** are identical in orientation; B₁₂ sticks coloured as C, ribbon colour; O, red; N, blue; P, orange. **d**, Superposition of ChCFeSP (grey), folate-free CFESP/MeTr (green), and folate-bound CFESP/MeTr

(orange) structures in **a**–**c**, highlighting one helix (thick ribbons) to show clamping motion (helix axes as straight blue lines) and B₁₂ (sticks) with 12-residue linker (thick ribbons) to B₁₂ domain (surface) to show swinging motion. **e**, Superposition of B₁₂ and CH₃-H₄folate in **d**, with Asn 199 shown for CFESP/MeTr structures in sticks (C, ribbon colour; O, red; N, blue). **f**, Same as **e**, with $2F_o - F_c$ density in blue (1.0σ) and pink mesh (4.0σ), and $F_o - F_c$ density in green mesh (3.0σ) for folate-bound CFESP/MeTr structure. Putative alternative B₁₂ corrin: cyan. **g**, Superposition of B₁₂ cofactors and CH₃-H₄folate in **f**.

the folate-free structure, B₁₂ in both folate-bound structures has indeed moved even closer to the MeTr folate-binding site (by an average of 7.7 Å) and exhibits new H-bonding features (Fig. 2c). In these folate-bound structures, the B₁₂ corrin ring has severed all interactions with the 'capping' CFESP small subunit and contacts only MeTr residues. Here, asparagine and glutamine residues that line the MeTr surface appear to participate in an 'amide hand off', sequentially passing B₁₂ along its trajectory as it progresses towards folate (Fig. 2b–c and Supplementary Fig. 10).

Interestingly, the terminal amide in this 'hand off', Asn 199, is strictly conserved in both MeTr and MethH and was previously shown to switch conformations between folate-free and folate-bound forms¹⁰, a feature also observed in the CFESP/MeTr structures presented here (Fig. 2b, c, e). In apo-MeTr structures, Asn 199 points upwards and out of the active site, whereas in folate-bound MeTr structures Asn 199 turns down to H-bond with the N⁵ of folate. Because N199A mutation moderately hinders folate binding (20-fold in dissociation constant, K_d) but dramatically compromises catalytic efficiency (k_{cat}/K_m) by 25,000-fold, Asn 199 is thought to be important for formation of the transition state¹⁰. In our CFESP/MeTr structures, we observe a new role for Asn 199 in B₁₂ domain conformational switching: when folate is absent, Asn 199 points out of the active site, blocking a closer B₁₂ position. However, when folate binds and Asn 199 reorients to H-bond with folate, space becomes available for B₁₂ to move closer to the MeTr folate-binding site. Therefore, the position of Asn 199 itself could help shift the conformational equilibrium of the B₁₂ domain, signalling that substrate has bound to MeTr. Asn 199 is an ideal signal for substrate binding, as it is the only MeTr residue known to reposition upon folate binding¹⁰.

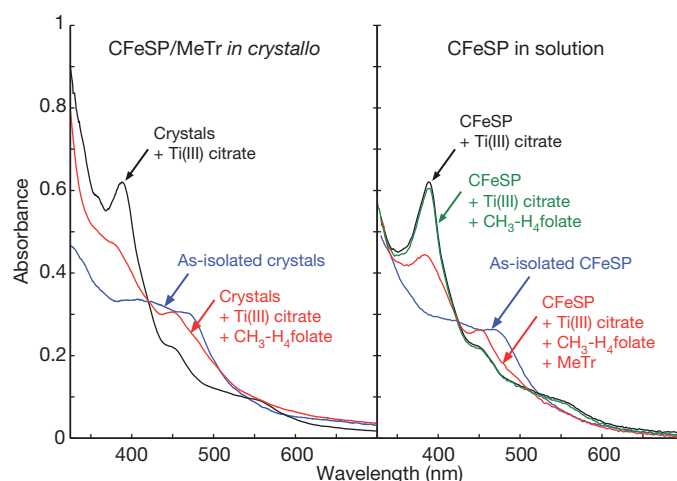


Figure 3 | Methyl transfer activity of CFESP/MeTr crystals by ultraviolet-visible absorption spectroscopy. As-isolated spectra (blue lines) for CFESP/MeTr crystals and CFESP in solution similarly have broad features at approximately 400 and 470 nm arising from the Fe₄S₄ cluster and the B₁₂ corrin. Following established protocols^{18,19,22,23,26,27}, B₁₂ reduction was achieved with Ti(III) citrate, yielding a sharp 390 nm peak indicative of active Co(I) in both solution and *in crystallo* spectra (black lines). Further treatment with CH₃-H₄folate yields decreased absorbance at 390 nm and a new peak at 450 nm (red lines), characteristic of the product complex (protein-bound CH₃-Co(III))^{18,19,22,23}. A control reaction (green line) confirms that turnover does not occur from free CH₃-H₄folate without MeTr, and the 450 nm peak indicates that B₁₂ remains CFESP-bound (free B₁₂ has a peak at approximately 520 nm instead^{18,21,27}).

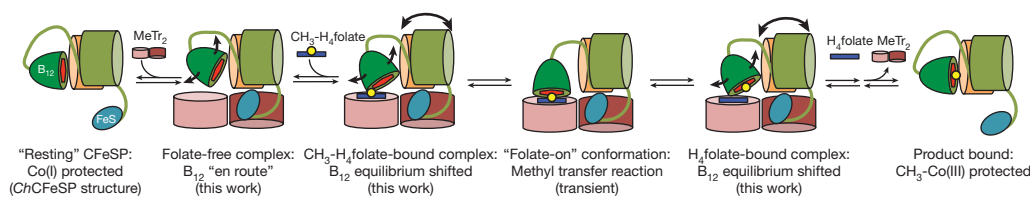


Figure 4 | Cartoon model of B₁₂-dependent methyl transfer in CFeSP/MeTr. For simplicity, only one of the two CFeSP heterodimers is shown. Protein domains are coloured as in Fig. 1, loops represent linkers, red hexagon

Displacement of the B₁₂ domain from its resting position to the position nearest the folate-binding site can be attributed to two independent conformational changes within the complex, best described as 'swinging' and 'clamping' motions (Fig. 2d). The B₁₂ domain can 'swing' relative to the rest of CFeSP (Supplementary Fig. 11), and CFeSP can 'clamp' the B₁₂ domain towards the MeTr active site (Fig. 2d and Supplementary Fig. 12). Despite varying degrees of 'clamping' over a range of approximately 14° across the structures, the interface between the CFeSP small subunit and MeTr is preserved (Supplementary Fig. 13).

Although folate binding shifts the average position of the B₁₂ domain closer to the MeTr folate-binding site, the B₁₂ Co is still too far for S_N2 methyl transfer (Fig. 2e). Intriguingly, a large, continuous electron density peak is present in 2F_o - F_c, F_o - F_c, and composite omit maps, emanating from the corrin ring and stretching directly over the folate-binding site, suggestive of an alternative, low-occupancy corrin conformation (Fig. 2f and Supplementary Fig. 14). A trial refinement of a putative corrin ring at 40% occupancy satisfies the F_o - F_c difference maps (Supplementary Fig. 15), positioning B₁₂ over the folate-binding site.

The multiple positions of B₁₂ captured here (Fig. 2g) highlight the conformational flexibility of the CFeSP/MeTr scaffold and provide a framework to understand the molecular juggling of domains during B₁₂-dependent methyl transfer (Fig. 4). Before MeTr binding, the CFeSP B₁₂ domain rests against the 'capping' small subunit, as in the structure of ChCFeSP⁹, with reactive Co(I) of B₁₂ protected ('resting' state). From this conformation, either the 'cap' or the B₁₂ domain must move to allow substrate access. Our folate-free CFeSP/MeTr structure indicates that upon MeTr binding, the B₁₂ domain becomes 'loosened' and flexible, adopting an ensemble of conformations that lie en route towards the MeTr active site. Here, the reactive B₁₂ species would be protected by the CFeSP small subunit and MeTr TIM barrels. CH₃-H₄folate binding to MeTr accompanied by movement of Asn 199 shifts the equilibrium of B₁₂ domain conformers, placing B₁₂ closer to folate, as in our folate-bound CFeSP/MeTr structures, with B₁₂ protection afforded by MeTr. It is notable that even after CH₃-H₄folate binds, the major B₁₂ position is still not directly over the folate methyl group, as such a position is expected to be transient. After methyl transfer, the B₁₂ domain can return to the small subunit to 're-cap' the methylated B₁₂ product, protected by the small subunit TIM barrel.

Overall, our data indicate that B₁₂ domain movement is not a simple, two-state switch between 'resting' and 'catalytic' conformations. Instead, a flexible B₁₂ domain samples an ensemble of conformations, where subtle shifts of the conformational equilibrium place B₁₂ progressively closer to the active site, thereby increasing the population of conformers capable of methyl transfer without obstructing substrate access or hindering domain movement. This model is consistent with MetH studies where ligation, alkylation and redox state of the B₁₂ cobalt can favour/disfavour various binding modes, alternately shifting the equilibrium of conformers for ordered domain rearrangements during the reaction cycle^{11,12,24,25}. We further identify MeTr residues that contact B₁₂ along its trajectory, ending with Asn 199. In MetH the B₁₂ ligating residue His 759 has been shown to play a dual role in catalysis and in signalling conformational shifts¹¹. The strictly conserved, folate-binding Asn 199 of MeTr could similarly play a dual role in both catalysis and conformational signalling. We thus expect this

is B₁₂, blue rectangle is folate and transferred methyl group is shown as a yellow sphere. Curved arrows denote 'swinging' and 'clamping' motions.

model for dynamic domain juggling, communicated by residues involved in substrate and cofactor binding, to be a common theme in methyl transfer between the B vitamins folate and B₁₂.

METHODS SUMMARY

CFeSP and MeTr were expressed and purified anaerobically from *M. thermoacetica* ATCC 39073 and from recombinant *Escherichia coli*, respectively. Crystals were grown anaerobically by hanging drop vapour diffusion. Diffraction data were collected at 24ID-C at the Advanced Photon Source, Argonne National Laboratory, and 5.0.2 and 8.2.2 at the Advanced Light Source, Lawrence Berkeley National Laboratory. Structures were solved by molecular replacement. Data collection and refinement statistics are presented in Supplementary Table 1, and representative electron densities for protein domains and for cofactors/substrate are shown in Supplementary Figs 16–21. Solution and *in crystallo* absorption spectra were collected as described in the text and in the Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 December 2010; accepted 3 February 2012.

Published online 14 March 2012.

- Matthews, R. G. Cobalamin-dependent methyltransferases. *Acc. Chem. Res.* **34**, 681–689 (2001).
- Banerjee, R. B. & Ragsdale, S. W. The many faces of vitamin B₁₂: catalysis by cobalamin-dependent enzymes. *Annu. Rev. Biochem.* **72**, 209–247 (2003).
- Matthews, R. G., Koutmos, M. & Datta, S. Cobalamin-dependent and cobamide-dependent methyltransferases. *Curr. Opin. Struct. Biol.* **18**, 658–666 (2008).
- Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G. & Ludwig, M. L. How a protein binds B₁₂: a 3.0 Å X-ray structure of B₁₂-binding domains of methionine synthase. *Science* **266**, 1669–1674 (1994).
- Dixon, M. M., Huang, S., Matthews, R. G. & Ludwig, M. L. The structure of the C-terminal domain of methionine synthase: presenting S-adenosylmethionine for reductive methylation of B₁₂. *Structure* **4**, 1263–1275 (1996).
- Doukov, T., Seravalli, J., Stezowski, J. J. & Ragsdale, S. W. Crystal structure of a methyltetrahydrofolate- and corrinoid-dependent methyltransferase. *Structure* **8**, 817–830 (2000).
- Bandarian, V. *et al.* Domain alternation switches B₁₂-dependent methionine synthase to the active conformation. *Nature Struct. Biol.* **9**, 53–56 (2002).
- Evans, J. C. *et al.* Structures of the N-terminal module imply large domain motions during catalysis by methionine synthase. *Proc. Natl Acad. Sci. USA* **101**, 3729–3736 (2004).
- Svetlitchnaia, T., Svetlitchnyi, V., Meyer, O. & Dobbek, H. Structural insights into methyltransfer reactions of a corrinoid iron-sulfur protein involved in acetyl-CoA synthesis. *Proc. Natl Acad. Sci. USA* **103**, 14331–14336 (2006).
- Doukov, T. I., Hemmi, H., Drennan, C. L. & Ragsdale, S. W. Structural and kinetic evidence for an extended hydrogen-bonding network in catalysis of methyl group transfer: role of an active site asparagine residue in activation of methyl transfer by methyltransferases. *J. Biol. Chem.* **282**, 6609–6618 (2007).
- Datta, S., Koutmos, M., Patridge, K. A., Ludwig, M. L. & Matthews, R. G. A disulfide-stabilized conformer of methionine synthase reveals an unexpected role for the histidine ligand of the cobalamin cofactor. *Proc. Natl Acad. Sci. USA* **105**, 4115–4120 (2008).
- Koutmos, M., Datta, S., Patridge, K. A., Smith, J. L. & Matthews, R. G. Insights into the reactivation of cobalamin-dependent methionine synthase. *Proc. Natl Acad. Sci. USA* **106**, 18527–18532 (2009).
- Banerjee, R. B. & Matthews, R. G. Cobalamin-dependent methionine synthase. *FASEB J.* **4**, 1450–1459 (1990).
- Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood–Ljungdahl pathway of CO₂ fixation. *Biochim. Biophys. Acta* **1784**, 1873–1898 (2008).
- Schrauzer, G. N. & Deutsch, E. Reactions of cobalt(II) superoxide nucleophiles. The alkylations of vitamin B₁₂s, cobaloximes(I), and related compounds. *J. Am. Chem. Soc.* **91**, 3341–3350 (1969).
- Harder, S. R., Lu, W.-P., Feinberg, B. A. & Ragsdale, S. W. Spectroelectrochemical studies of the corrinoid/iron-sulfur protein involved in acetyl coenzyme A synthesis by *Clostridium thermoaceticum*. *Biochemistry* **28**, 9080–9087 (1989).
- Banerjee, R. B., Harder, S. R., Ragsdale, S. W. & Matthews, R. G. Mechanism of reductive activation of cobalamin-dependent methionine synthase: an electron

- paramagnetic resonance spectroelectrochemical study. *Biochemistry* **29**, 1129–1135 (1990).
18. Menon, S. & Ragsdale, S. W. Role of the [4Fe-4S] cluster in reductive activation of the cobalt center of the corrinoid iron–sulfur protein from *Clostridium thermoaceticum* during acetate biosynthesis. *Biochemistry* **37**, 5689–5698 (1998).
 19. Menon, S. & Ragsdale, S. W. The role of an iron–sulfur cluster in an enzymatic methylation reaction. *J. Biol. Chem.* **274**, 11513–11518 (1999).
 20. Zydowsky, T. M. *et al.* Stereochemical analysis of the methyl transfer catalyzed by cobalamin-dependent methionine synthase from *Escherichia coli* B. *J. Am. Chem. Soc.* **108**, 3152–3153 (1986).
 21. Ragsdale, S. W., Lindahl, P. A. & Münck, E. Mössbauer, EPR, and optical studies of the corrinoid/iron–sulfur protein involved in the synthesis of acetyl coenzyme A by *Clostridium thermoaceticum*. *J. Biol. Chem.* **262**, 14289–14297 (1987).
 22. Zhao, S., Roberts, D. L. & Ragsdale, S. W. Mechanistic studies of the methyltransferase from *Clostridium thermoaceticum*: origin of the pH dependence of the methyl group transfer from methyltetrahydrofolate to the corrinoid/iron–sulfur protein. *Biochemistry* **34**, 15075–15083 (1995).
 23. Seravalli, J., Zhao, S. & Ragsdale, S. W. Mechanism of transfer of the methyl group from (6S)-methyltetrahydrofolate to the corrinoid/iron–sulfur protein catalyzed by the methyltransferase from *Clostridium thermoaceticum*: a key step in the Wood–Ljungdahl pathway of acetyl–CoA synthesis. *Biochemistry* **38**, 5728–5735 (1999).
 24. Jarrett, J. T. *et al.* Mutations in the B₁₂-binding region of methionine synthase: how the protein controls methylcobalamin reactivity. *Biochemistry* **35**, 2464–2475 (1996).
 25. Bandarian, V., Ludwig, M. L. & Matthews, R. G. Factors modulating conformational equilibria in large modular proteins: a case study with cobalamin-dependent methionine synthase. *Proc. Natl Acad. Sci. USA* **100**, 8156–8163 (2003).
 26. Wirt, M. D. *et al.* Structural and electronic factors in heterolytic cleavage: formation of the Co(I) intermediate in the corrinoid/iron–sulfur protein from *Clostridium thermoaceticum*. *Biochemistry* **34**, 5269–5273 (1995).
 27. Stich, T. A. *et al.* Spectroscopic studies of the corrinoid/iron–sulfur protein from *Moorella thermoacetica*. *J. Am. Chem. Soc.* **128**, 5010–5020 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. E. Darty for his assistance with the purification of CFeSP. This work was supported by National Institutes of Health grants GM69857 (to C.L.D.) and GM39451 (to S.W.R.) and the MIT Energy Initiative (to C.L.D.). C.L.D. is a Howard Hughes Medical Institute Investigator. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the US Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract No. DE-AC02-05CH11231.

Author Contributions Y.K. performed crystallization and data collection, processing and refinement that gave the folate-free and folate-bound CFeSP/MeTr structures. N.A. built the microspectrophotometer and performed *in crystallo* spectroscopic experiments with the aid of Y.K., who performed the parallel solution spectroscopic experiments. T.I.D. determined initial crystallization conditions and performed initial data collection, and L.C.B. processed and refined these data. G.B. and J.S. expressed and purified protein samples, and S.W.R. and C.L.D. were involved in study design. Y.K. and C.L.D. wrote the manuscript.

Author Information Atomic coordinates are deposited in the Protein Data Bank under accession codes 4DJD, 4DJE and 4DJF. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.L.D. (cdrennan@mit.edu).

METHODS

Protein purification. CFeSP was expressed and purified anaerobically from *M. thermoacetica* ATCC 39073 as described²¹, except for the following modifications. CFeSP was purified from cell extracts using DEAE-cellulose and high-resolution Q-Sepharose anion exchange chromatography followed by phenyl-Sepharose hydrophobic interaction chromatography. Fractions containing CFeSP were concentrated and buffer exchanged using Amicon ultracentrifuge concentrators in the anaerobic chamber. MeTr was expressed and purified anaerobically from recombinant *E. coli* as described⁶. Concentrations of CFeSP and MeTr protein samples were determined using the Rose-Bengal method²⁸ and kept in storage buffer: 50 mM Tris-HCl, pH 7.6, 100 mM NaCl, 2 mM dithiothreitol.

Crystallization. Crystals of the folate-free CFeSP/MeTr complex were grown by hanging-drop vapour diffusion in an anaerobic chamber (Coy Laboratories) at room temperature by adding 1 µl of precipitant (100 mM Bis-Tris, pH 6.5, 100 mM calcium acetate, 9% PEG 5,000 monomethyl ether, 20% glycerol) to 2 µl of an equimolar mixture of CFeSP and MeTr (approximately 250 µM monomer for each), over a 0.5 ml reservoir solution of precipitant. Large, brown, rod-shaped crystals appeared overnight. Crystals were looped and cryo-cooled in liquid nitrogen anaerobically before collection of X-ray diffraction data at 100 K. Crystals of CFeSP/MeTr co-crystallized with the CH₃-H₄folate substrate were obtained in the same manner as above, except the protein solution also contained CH₃-H₄folate at 1 mM concentration. Crystals of CFeSP/MeTr co-crystallized with both CH₃-H₄folate and Ti(III) citrate as a reductant were obtained in the same manner, except the precipitant solution also contained Ti(III) citrate at 3 mM concentration.

Structure determination of folate-free CFeSP/MeTr structure. Two X-ray diffraction data sets were collected for the folate-free CFeSP/MeTr structure. A lower-resolution data set (3.3 Å) was collected at the Advanced Light Source (ALS) beam line 5.0.2 ($\lambda = 1.1000$ Å), and a higher-resolution data set (2.38 Å) was later collected at the Advanced Photon Source (APS) beam line 24ID-C ($\lambda = 0.9792$ Å).

The initial data set to 3.3 Å resolution was processed in HKL2000 and Scalepack²⁹. The structure was solved by molecular replacement in Phaser³⁰, using individual structures of MeTr¹⁰ (Protein Data Bank accession number 2E7F) and ChCFeSP⁹ (Protein Data Bank accession number 2H9A) lacking its B₁₂ domain as independent search models. Two CFeSP/MeTr complexes (approximately 220 kDa each) were found in the asymmetric unit, and crystals belonged to the space group *P*₂₁₂₁ with unit cell dimensions (Å): *a* = 137.42, *b* = 159.87 and *c* = 241.92. Iterative rounds of refinement with residue-grouped *B*-factors were performed in CNS³¹ and PHENIX³², with model building in Coot³³. The four B₁₂ domains present in the asymmetric unit were kept as a polyaniline model. Final *R*-factors for working and test reflections (*R*_{work} and *R*_{free}) were 29.2% and 33.7%, respectively, when refinement of the structure to higher resolution began. Data collection and refinement statistics for this data set are shown in Supplementary Table 1. Ramachandran analysis was performed in PROCHECK³⁴: 74.0% of residues resided in the most favoured region, with 21.2% additionally allowed, 3.1% generously allowed and 1.7% disfavoured.

The data set to 2.38 Å resolution was processed in HKL2000 and Scalepack²⁹. Although this crystal formed in similar conditions as the crystal which gave the 3.3 Å resolution data set, the space group was now *P*₂₁₂, with unit cell dimensions (Å): *a* = 125.71, *b* = 242.84 and *c* = 79.67. The structure for this crystal was thus solved by molecular replacement in Phaser³⁰ using the MeTr homodimer and CFeSP heterodimers lacking B₁₂ domains from the previously refined model of the 3.3 Å resolution structure as independent search models. Only one CFeSP/MeTr complex was present in the asymmetric unit. Iterative rounds of refinement were performed in CNS³¹ and PHENIX³², with model building in Coot³³. Translation/libration/screw refinement was performed in latter refinement rounds with seven translation/libration/screw groups: the MeTr homodimer (chains A and B), the Fe₄S₄ domain of one CFeSP large subunit (chain C), the TIM domain of chain C with one small subunit (chain D), the B₁₂ domain of chain C, the Fe₄S₄ domain of the second large subunit (chain E), the TIM domain of chain E with the second small subunit (chain F), and the B₁₂ domain of chain E. Data collection and refinement statistics are shown in Supplementary Table 1, and average *B*-factors for each domain of the final model are given in Supplementary Table 2. Ramachandran analysis was performed in PROCHECK³⁴: 90.1% of residues resided in the most favoured region, with 9.5% additionally allowed, 0.3% generously allowed and 0.2% disfavoured. The final model contains residues 1–262 (of 262) for both MeTr chains (A and B), residues 2–442 (of 446) for both CFeSP large subunit chains (C and E) and residues 1–323 (of 323) for both CFeSP small subunit chains (D and F).

Except for the Fe₄S₄ and B₁₂ domains, the entire structure is composed of TIM barrels for which the electron density is well-defined (Supplementary Fig. 16). Electron density is weaker for the Fe₄S₄ domains (Supplementary Figs 5, 17 and

19), consistent with the fact that these domains exhibit higher *B*-factors (Supplementary Table 2 and Supplementary Fig. 7). However, reasonable electron density is present for the main chain and most side chains, allowing us to build a model for this domain. Still, several side chains of the Fe₄S₄ domains lack clear electron density; thus, for these residues, atoms were truncated past the Cβ atom (chain C, 12 residues truncated and chain E, 16 residues truncated, out of 56 total residues).

Although *B*-factors are high and electron density is weak for the B₁₂ domains in general, electron density for the B₁₂ cofactors is unambiguous (Supplementary Fig. 20), and density is also clear in several helical regions, including those near B₁₂ (Supplementary Fig. 18). Because the structure of a CFeSP B₁₂ domain bound with B₁₂ was already known⁹, we used the clear electron density of the B₁₂ cofactor and the resolvable helices to position the B₁₂ domain during model building. Still, many side chains of the CFeSP B₁₂ domains lacked clear electron density, and thus for these residues, atoms were truncated past the Cβ atom (chain C, 50 residues truncated and chain E, 59 residues truncated, out of 118 total residues).

The B₁₂ cofactor in the final model contains 5,6-dimethylbenzimidazole as the lower ligand moiety, as in cobalamin. Although active with cobalamin³⁵, previous studies have shown that CFeSP isolated from *M. thermoacetica* harbours an unusual B₁₂ derivative that contains 5'-methoxybenzimidazole as the lower ligand instead³⁶. However, disorder of the B₁₂ cofactor and B₁₂ domain owing to thermal motion of these regions in the CFeSP/MeTr crystal resulted in weak electron density for substituents of the benzimidazolyl ring (Supplementary Figs 18 and 20). Therefore, we cannot confirm the presence of this unusual B₁₂ derivative from our crystallographic studies, and we have thus modelled cobalamin as the form of B₁₂ in the structure.

Previous spectroscopic studies²⁷ in addition to the crystal structure of ChCFeSP⁹ have indicated that a water molecule coordinates the central cobalt of B₁₂ in the as-isolated CFeSP. Here, Co(II) is the major species and is expected to be five-coordinate. However, because of disorder we do not observe electron density to suggest a water molecule bound to cobalt (Supplementary Figs 18 and 20). Accordingly, we have not modelled a water molecule.

Structure determination of folate-bound CFeSP/MeTr structures. For crystals grown with CH₃-H₄folate, X-ray diffraction data were collected at APS beam line 24ID-C to 3.50 Å resolution at $\lambda = 1.6039$ Å to optimize the cobalt peak anomalous signal. For crystals grown with both CH₃-H₄folate and Ti(III) citrate, X-ray diffraction data were collected at ALS beam line 8.2.2 to 3.03 Å resolution at $\lambda = 1.0000$ Å. The structures were solved by molecular replacement using the MeTr homodimer and CFeSP heterodimers lacking Fe₄S₄ and B₁₂ domains from the folate-free 2.38 Å CFeSP/MeTr structure as independent search models. Refinement of the folate-free CFeSP/MeTr structure against either folate-bound X-ray data set was not sufficient to solve the structure, as the unit cell dimensions were markedly different (Supplementary Table 1). After molecular replacement, one CFeSP/MeTr complex was present in the asymmetric unit, and omit electron density clearly indicated the presence of bound folate (Supplementary Fig. 21). Iterative rounds of refinement were performed in CNS³¹ and PHENIX³², with model building in Coot³³. The same test set of reflections for *R*_{free} calculations was used for both folate-bound data sets. Data collection and refinement statistics are shown in Supplementary Table 1. Ramachandran analysis was performed in PROCHECK³⁴: for the CH₃-H₄folate-only structure, 89.6% of residues resided in the most favoured region, with 9.8% additionally allowed, 0.3% generously allowed and 0.2% disfavoured. For the CH₃-H₄folate with Ti(III) citrate structure, 89.5% of residues resided in the most favoured region, with 10.0% additionally allowed, 0.2% generously allowed and 0.3% disfavoured. The final models both contain folate, B₁₂ and residues 1–262 (of 262) for MeTr chains (A and B), residues 2–442 (of 446) for CFeSP large subunit chains (C and E) and residues 1–323 (of 323) for CFeSP small subunit chains (D and F). As with the folate-free structure, several side chains of the Fe₄S₄ domains for both folate-bound structures lacked clear electron density; thus for these residues, atoms were truncated past the Cβ atom (chain C, 15 residues truncated and chain E, 18 residues truncated, out of 52 total residues). Similarly, many side chains of the B₁₂ domains lacked electron density, and thus for these residues, atoms were truncated past the Cβ atom (chain C, 51 residues truncated and chain E, 78 residues truncated, out of 118 total residues). The liganded/oxidation states of folate and B₁₂ in these structures were determined by use of a microspectrophotometer (see below).

Solution and in crystallo ultraviolet-visible absorption spectroscopy to determine enzyme activity in crystallo. Titanium(III) citrate (100 mM in 50 mM Tris, pH 7.6) was prepared³⁷, and (6S)-5-methyl-5,6,7,8-tetrahydrofolate (CH₃-H₄folate) containing one glutamate tail was purchased from Schircks Laboratories. As-isolated, reduced and methylated CFeSP samples in solution were prepared in a room-temperature anaerobic chamber (MBraun) following similar procedures to those previously described^{18,19,22,23,26,27}. Briefly, purified CFeSP (20 µM) was used for the as-isolated sample, CFeSP mixed with Ti(III) citrate (1 mM) was used for

the reduced sample, and CFeSP mixed with equimolar MeTr, Ti(III) citrate (1 mM), and CH₃-H₄folate (1 mM) was used for the methylated sample. Spectra were taken using a Nanodrop 2000c (Thermo Scientific) in a quartz cuvette or on the sample stage in the anaerobic chamber directly after mixing; identical solutions lacking CFeSP were used as blanks.

To obtain *in crystallo* absorption spectra, CFeSP/MeTr crystals in as-isolated, reduced and methylated forms were prepared in a similar fashion. In a room-temperature anaerobic chamber (Coy Laboratories), crystals were looped into a 2 µl drop, which was placed on a cover slide and contained one of the following three solutions for as-isolated, reduced and methylated samples, respectively: well solution, well solution with Ti(III) citrate (10 mM) and well solution with Ti(III) citrate (10 mM) and CH₃-H₄folate (1 mM). A ring of epoxy surrounding each drop was applied to the cover side, and a second cover slide was placed on top, sandwiching the drops within a uniform distance separation and sealing the crystals within an anaerobic environment. Upon curing of the epoxy, crystals were brought out of the anaerobic chamber and mounted on an XZ translation stage (Newport, UMR8.25 and SM-13) in a fibre optic coupled microspectrophotometer (Ocean Optics, Jaz) with 40 mm diameter reflective objectives (Optique Peter) and a deuterium-halogen lamp (DH2000-BAL, Ocean Optics), similar to that previously described^{38,39}. Stray light was blocked with blackout material. The light focus was coarsely aligned to the crystals by visual inspection and then finely aligned by monitoring light transmission in real time. Data were acquired at room temperature with the SpectraSuite software (Ocean Optics). The background transmission was measured through the solution immediately surrounding the crystals. The dark current was measured with the light shuttered off. Sample, reference and dark current spectra were acquired by averaging 10–50 scans with total exposure times of 90–1000 ms. Experiments were completed within 60 min of sample preparation, and crystals remained intact over the course of the experiment, as observed using a microscope after data collection.

To generate Fig. 3, absorbance spectra were scaled relative to each other to account for variable crystal sizes and path lengths, where absolute peak absorbances did not exceed one absorbance unit.

***In crystallo* ultraviolet–visible absorption spectroscopy on folate-bound CFeSP/MeTr crystals to determine liganded/oxidation state of bound B₁₂ and folate.** Ultraviolet–visible absorption spectra were collected on a microspectrophotometer at 100 K for crystals of folate-free CFeSP/MeTr, crystals that were grown in the presence of CH₃-H₄folate only and for crystals grown in the presence of both CH₃-H₄folate and Ti(III) citrate (Supplementary Fig. 9). The spectra were compared with the analogous solution spectra (Fig. 3). The spectrum for the folate-free crystal was similar to the spectrum of CFeSP alone in solution,

with broad features at approximately 400 and 470 nm indicative of the Fe₄S₄ cluster and B₁₂ primarily in the Co(II) state. The spectrum for the crystal grown with CH₃-H₄folate matched the spectrum of folate-free crystals, indicating that B₁₂ had remained primarily in the Co(II) state, and turnover had not occurred. However, the spectrum for the crystal grown with both CH₃-H₄folate and Ti(III) citrate was markedly different and contained a peak at 450 nm, indicating that B₁₂ in these crystals was methylated to the CH₃-Co(III) state. Based on these data, we modelled the methyl group on folate in the structure co-crystallized with CH₃-H₄folate only, whereas we modelled the methyl group bound to Co of B₁₂ for the CH₃-H₄folate/Ti(III) citrate structure. Without these spectroscopic data, assignment of the location of the methyl group would otherwise have been prevented by the resolution limits of the data.

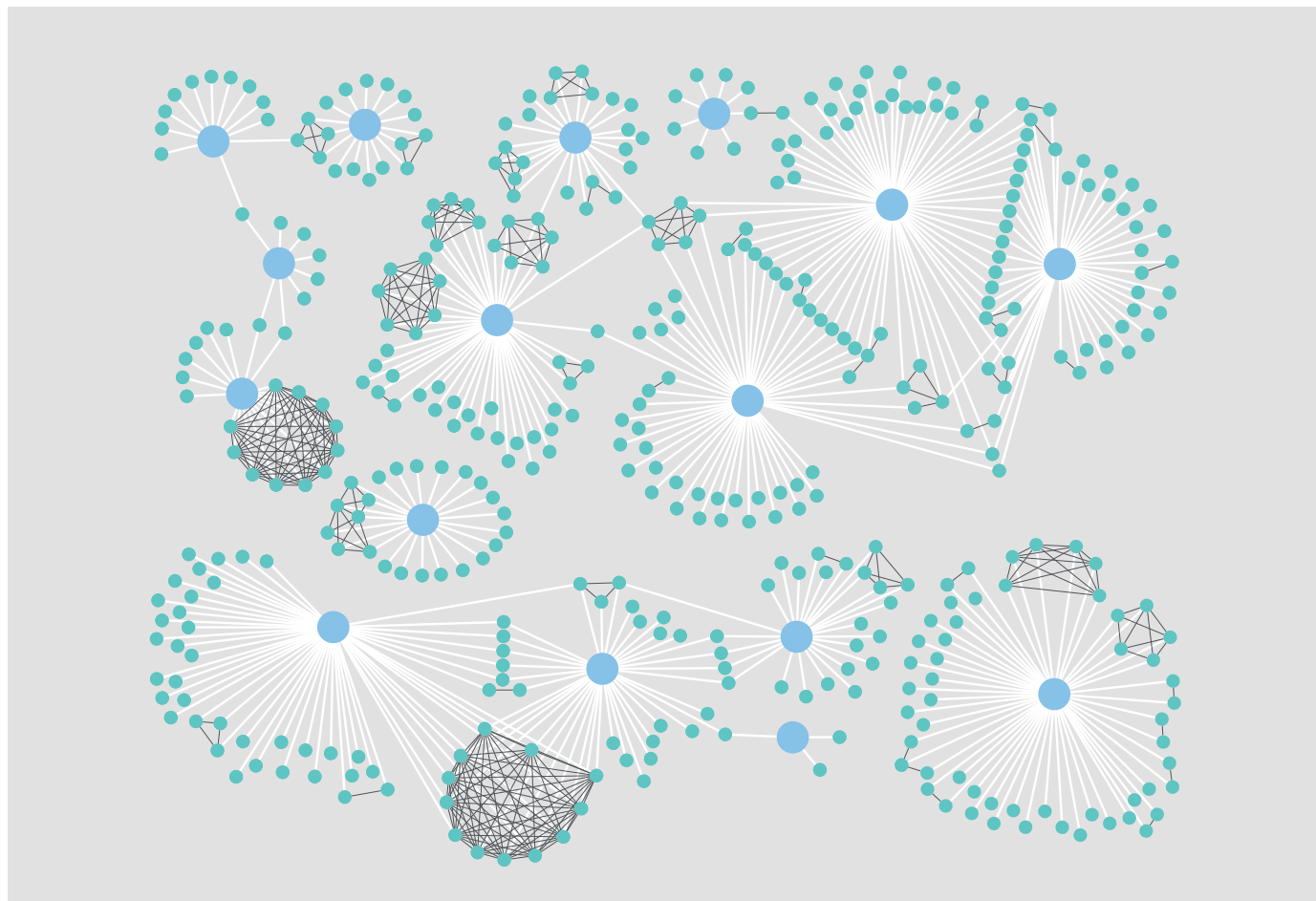
28. Elliott, J. I. & Brewer, J. M. The inactivation of yeast enolase by 2,3-butanedione. *Arch. Biochem. Biophys.* **190**, 351–357 (1978).
29. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
30. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
31. Brünger, A. T. *et al.* Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
32. Adams, P. D. *et al.* PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
33. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
34. Laskowski, R. A., MacArthur, M. W. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
35. Lu, W.-P., Schiau, I., Cunningham, J. R. & Ragsdale, S. W. Sequence and expression of the gene encoding the corrinoid/iron–sulfur protein from *Clostridium thermoaceticum* and reconstitution of the recombinant protein to full activity. *J. Biol. Chem.* **268**, 5605–5614 (1993).
36. Jungdahl, L. G., LeGall, J. & Lee, J.-P. Isolation of a protein containing tightly bound 5-methoxybenzimidazolylcobamide (factor III_m) from *Clostridium thermoaceticum*. *Biochemistry* **12**, 1802–1808 (1973).
37. Zehnder, A. J. B. & Wuhrmann, K. Titanium(III) citrate as a nontoxic oxidation–reduction buffering system for the culture of obligate anaerobes. *Science* **194**, 1165–1166 (1976).
38. Royant, A. *et al.* Advances in spectroscopic methods for biological crystals. 1. fluorescence lifetime measurements. *J. Appl. Cryst.* **40**, 1105–1112 (2007).
39. Barstow, B., Ando, N., Kim, C. U. & Gruner, S. M. Alteration of citrine structure by hydrostatic pressure explains the accompanying spectral shift. *Proc. Natl Acad. Sci. USA* **105**, 13362–13366 (2008).

TECHNOLOGY FEATURE

THE INTERACTION MAP

As increasing numbers of protein–protein interactions are identified, researchers are finding ways to interrogate these data and understand the interactions in a relevant context.

M. SHALES/KROGAN LAB/UCSF



Multiple replicated experiments and sophisticated statistics reveal 497 interactions between 16 HIV proteins (blue) and hundreds of human factors.

BY MONYA BAKER

Around the time that scientists celebrated the completion of the draft sequence of the human genome, papers from two separate groups described results of another project that tested all the possible pairings of thousands of yeast proteins to see whether they interact^{1,2}.

The importance of protein–protein interactions is beyond dispute. Little happens in a cell without one protein ‘touching’ another. Whether a cell divides, secretes a hormone or triggers its own death, protein–protein interactions make the event happen. Consequently, comprehensive maps

showing which proteins came together in a yeast cell were much anticipated.

But the results took scientists aback. Although the two research groups had explored the full collection of proteins in the same organism using the same yeast two-hybrid (Y2H) assay, the two papers found fewer than 150 interactions in common — only 10% of the findings that either team dubbed high quality. Most scientists regarded the results as so riddled with artefacts that they were useless.

“As you can imagine, people were extremely critical. They just couldn’t believe that you would get such different results when you were studying the same thing,” recalls Peter

Uetz, who studies protein interactions at the Center for the Study of Biological Complexity at Virginia Commonwealth University in Richmond, and was a co-author on one of the papers¹. Even today, many researchers look askance at the Y2H assays used in the studies.

But Marc Vidal, a systems biologist at the Dana-Farber Cancer Institute in Boston, Massachusetts, says that the technique has come a long way in a decade. Not only have researchers found ways to recognize and reduce false-positives, but gruelling follow-up studies show that the startlingly low overlap between the two reports was not because the assays found so many interactions that do not exist, but because they missed so many that do³. ►

► Understanding these interactions is as important as ever. Protein interactomes — maps of protein interactions — are raw fuel for systems biologists. Promising techniques to block protein–protein interactions in cancer cells and for other diseases have launched a string of biotechnology deals. Considering disease in terms of protein–protein interactions rather than individual genes and proteins could help to untangle jumbled observations. For example, mutations in the same protein could lead to different diseases by disrupting different interactions. Similarly, mutations in different proteins that disrupt the same interaction could lead to the same disease.

A good reference map of interactions would be like completing the human genome sequence, says Vidal, and could spawn further efforts to study genetic variation and function. A validated network would give scientists a jumping off point for more experiments. “My guess is that as these networks grow, we will get more elaborate ways of understanding where these interactions take place, when and why,” he says. “We are getting a sense of a cell’s organizational self by doing this.”

SCREENING SYSTEMS

First described in 1989, the Y2H assay tests the interactivity of pairs of proteins by attaching them to two halves of a transcription factor⁴. If the proteins come together, the transcription factor is reformed, activating reporter genes and allowing the yeast to grow. Companies including Hybrigenics in Paris and Dualsystems Biotech in Zurich, Switzerland, run Y2H as a service.

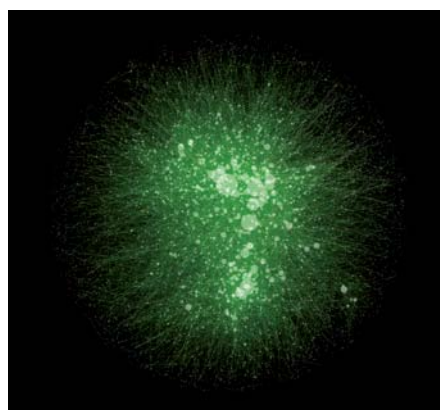
“Yeast two-hybrid has an enormous advantage, which might also be a disadvantage: it can detect low affinity,” says Erich Wanker, a neuroproteomics researcher at the Max Delbrück Center for Molecular Medicine in Berlin, and co-editor of a book on the topic⁵. In other words, the assay can identify weak, transient pairings such as those that perpetuate cell signalling. But it also detects proteins that randomly bump together. This bumping has led to almost philosophical discussions. “At what point do we really believe that it’s an interaction?” asks Wanker.

Scientists have also found ways to detect and avoid many sorts of false positive. Artefacts from ‘sticky’ proteins, which bind non-specifically to other proteins, can be identified and excluded. Growth that is promoted by a single introduced protein rather than a reformed transcription factor can also be recognized.

Precise systems also exist to make sure that all desired combinations are tested. Rather than transfecting the same yeast cell with genes for two potential interaction partners, yeast are transfected with individual genes, mated in pools and their progeny assayed for growth. Robotic systems mix yeast precisely and run multiple replicates of each assay. The

number of times that the same interaction is seen becomes part of a quality score. “Our view is that Y2H can give reliable and reproducible results,” says Wanker.

Still, some interactions will not be observed in Y2H. For example, the interacting proteins have to allow the two halves of the transcription factor to reunite, and the proteins must be able to reach the nucleus to activate the reporter gene. Thus, interactions



Interaction maps can help to explain protein function and identify new ways to fight disease.

with membrane- or organelle-specific proteins are invisible.

Besides Y2H, lower-throughput tests in mammalian cells can be used to screen interactions; these tests include luminescence-based mammalian interactome (LUMIER), mammalian protein–protein interaction trap (MAPPIT), protein arrays and protein-fragment complementation assay (PCA). Although these are orders of magnitude slower than Y2H, they can probe interactions in a more relevant context.

MAPPIT is one of the highest-throughput mammalian screens. Instead of a yeast transcription factor, a mammalian cytokine receptor is split and becomes capable of cell signalling only when reconstituted. In 2009, Jan Tavernier, a network biologist at VIB, a life-sciences research institute in Ghent, Belgium, described a higher throughput version of MAPPIT in which plasmids encoding potential interaction partners linked to one cytokine receptor fragment can be individually spotted into wells and stored⁶. To begin the experiment, wells are filled with cells expressing the cytokine-receptor fragment linked with the selected ‘bait’. When interactions occur, signalling activates the light-emitting enzyme luciferase.

Using multiwell plates it costs about €2,000 (US\$2,600) to screen one bait protein against the human ORFeome (a complete set of cloned protein-encoding open reading frames), says Tavernier, who hopes to describe techniques to run MAPPIT on microchips later this year. Miniaturized assays should reduce the cost to €100 and allow the ORFeome to be tested

against 100 baits a week.

At this throughput and cost, Tavernier says, new kinds of experiments become feasible. Instead of restricting screens to yeast cells, “you start mapping full interactomes in the appropriate species”, he says. In addition, Tavernier plans to compare how interactomes change when cells are treated with agents such as drugs or toxic chemicals. He is hoping to commercialize the technology, and is working with Vidal and other scientists to map human protein interactions using both MAPPIT and Y2H assays.

LUMIER assays are also relatively high-throughput and can be used to test whether particular interactions are affected by drugs, hormones or other additives. For these assays, cells are transiently transfected with two proteins. One protein is attached to a hydrophilic peptide called FLAG. Potential interaction partners are linked with luciferase. Cells are lysed, the FLAG-tagged proteins are captured and the presence of the interacting partners can be detected by the light they give off⁷.

Protein-fragment complementation assays, which can be conducted in yeast as well as mammalian cells, rely on reconstituting a wide range of ‘reporters’, often enzymes or fluorescent proteins. Since the reporters can signal throughout the cell, interactions can be detected where they naturally occur.

In a collection of articles published in January 2009, Vidal, Wanker and others described what Vidal terms an empirical framework for assessing protein interactions found in high-throughput screens³. In practice, this means repeating experiments using different types of assay and comparing the results with sets of controls. The positive

controls are a reference set of about 100 well-established interactions carefully selected from the literature. The negative controls are some 100 randomly assigned pairs that have never been observed together. Conditions of the assays are adjusted to boost detection of positive controls without raising the detection of random interactions.

As part of a framework put forth in

*Nature Methods*⁸, results from interaction studies should be confirmed in different types of assays. The more methods that find an interaction, the more confident researchers can be. Still, collectively, these assays detect only about 70% of the positive reference set (see ‘Beyond binary interactions’).



“We are getting a sense of a cell’s organizational self by having a validation network”

Marc Vidal



Yeast two-hybrid assays can probe hundreds of thousands of potential protein interaction pairs a week.

High-throughput experiments are not the only way to identify protein–protein interactions. Several databases, such as the Biological General Repository for Interaction Datasets (see thebiogrid.org) and IntAct (see www.ebi.ac.uk/intact), compile lists of interactions as they are published in the literature, culling from both small-scale

and high-throughput experiments as well as predicted interactions inferred from other analyses. But this list is not even close to complete, says Sandra Orchard, a proteomics service coordinator at the European Bioinformatics Institute in Hinxton, UK, who helped to develop minimal information standards to help share and evaluate interaction data. “We

will be lucky if as much as 30% of the yeast interactome has been observed,” she says. For the human interactome, she estimates that the figure is less than 10%, including published results that are not captured in the databases.

P. SONNABEND/MDC

WHEN TO BELIEVE

Biologists rely on interaction data in several ways. They often layer protein–protein interaction networks onto other networks. After identifying transcription factors that regulate a gene, for example, they search databases and literature for transcription factors’ interaction partners. Researchers also explore how sets of proteins are connected to each other, and then ask questions based on the structure of the network, such as classifying the proteins that have the most interaction partners. But not all interaction data are equal, warns Russell Finley, a network biologist at Wayne State University School of Medicine in Detroit, Michigan, who believes that incorporating quality measures could make the data substantially more powerful. At present, he says, savvy researchers filter out interactions unless they have been observed more than once through different methods, but these ‘intuitive filters’ can be biased. For example, the more often a protein is studied,

Beyond binary interactions

Rather than individual proteins, protein complexes are the functioning biochemical entities in a cell, says Nevan Krogan, a systems biologist at the University of California, San Francisco. “If you think about a protein, more often than not it is in a protein complex.” If a complex requires multiple components to form, two-hybrid studies cannot be expected to detect it.

Instead, complexes are generally studied in ‘pull down’ assays. The gene for a protein of interest is fused to a peptide tag that allows it to be ‘fished’ from cell lysates; less commonly, antibodies are used to purify out unlabelled proteins. The captured protein pulls its associates with it, and these are identified, usually by mass spectrometry. Companies such as Agilent Technologies in Santa Clara, California; Cell Signaling Technology in Danvers, Massachusetts; and IBA in Göttingen, Germany, offer transfection vectors and capture technologies.

Typically, researchers tag one protein of interest in each experiment, and these are then captured on a column. If not enough protein is captured, researchers can use experiments that add further tags and capture steps. If a complex is fragile, crosslinking reagents can be added to cells, binding nearby proteins together.

Although transient and fragile components

of complexes are hard to detect, the biggest problem in pull down assays is background — proteins that are not part of a complex but get pulled along. Researchers are getting much better at telling which proteins observed with a tagged protein are actually part of a complex, says Anne-Claude Gavin, who studies protein complexes at the European Molecular Biology Laboratory in Heidelberg, Germany.

The secret is statistics. If a protein complex has ten components, there should be ten ways to pull down the complex, explains Gavin. In a large data set involving many tagged proteins, she says, the same protein should occur in three forms: as the tagged protein, as an interaction partner and as background. Several scoring systems are used to sort artefacts from real interactions and to identify the components of protein complexes. The key, says Krogan, is to collect enough data. “For these scoring systems to be effective, you need pull downs of many types of proteins.”

Krogan and his colleagues¹² recently used pull down experiments to identify complexes formed between HIV proteins and host proteins. Disrupting these interactions could prevent the virus from entering cells. They selected 18 HIV proteins and tagged them in two ways, using FLAG and Strep tags. The

team developed a scoring system called mass spectrometry interaction statistics to sort the interactions. The system compiles a single score on the basis of abundance (more abundant proteins are more likely to be background), reproducibility and specificity. Almost 500 proteins reached this threshold; only 19 of these had already been reported¹².

This system vastly increased the number of potential targets for drugs that could treat HIV, and revealed a way to perform further studies on a previously intractable HIV protein — Vif, which was known to overcome human cells’ antiviral defences but was hard to study biochemically and structurally. The reason, says Krogan, was missing interaction partners. “You needed all the components of the complex there to make it behave.”

Now, researchers will be able to find potential ways to disrupt the complex, which could lead to new anti-HIV drugs. The real challenge, says Krogan, is to continue to mine the identified interaction partners for biological meaning: low-throughput work that requires extensive follow-up study. “The main goal of making these maps is not generating these maps. It is to extract biological insight, mechanism and testable hypotheses,” he says. “Sadly, work almost always stops at the maps.” **M.B.**

the more interactions will be found. Finley says that a better approach would be to consider all the data available and assign a score reflecting the likelihood that an interaction is real. Computer analyses could then be used to consider more interactions, giving more weight to those with higher confidence scores.

But an interaction can occur and have no actual consequences. "The real question is what interactions have meaning in the first place," says Stephen Michnick, a biochemist at the University of Montreal, Canada. "An interaction can be quite good, that is, reproducible in multiple assays, but not be biologically important." In other words, the interaction has no discernible effects: it does not start or stop a molecular machine, activate an enzyme or send another protein to destruction.

Michnick came to these conclusions after conducting a comprehensive study that allowed protein interactions to be studied in a more natural context. In the protein-fragment complementation assays, interacting proteins reconstituted an enzyme that yeast needed to survive under culture conditions⁹. This identified about 3,000 new interactions, with many involving membrane and other proteins that cannot reach the cell nucleus.

But thousands of other protein interactions were observed with less confidence. "We were surprised that there were known proteins that made too many interactions or made interactions that didn't make biological sense," Michnick recalls. "We thought we had the perfect method, and so we would get perfect results." "So we thought, if we are seeing junk interactions and other people are seeing junk, what is the junk?" The answer, he believes, is that these are naturally occurring 'junk' interactions that, like sections of DNA that do not seem to have a function, simply exist.

Michnick believes that perhaps as many as half of the interactions observed even in rigorous screens have no biological function. Abundant proteins should be treated with particular scepticism, but if the same pairs of proteins are consistently found together and not with other proteins then that interaction is more likely to be real, and the same is true of interactions identified across multiple species. "The parts that are functional have to be dissected from the rest of what's there," Michnick says.

Trey Ideker, a network biologist at the University of California, San Diego, is more worried that such a small percentage has been observed at all. "It's not clear how you can shortcut to the functional interactions without some unbiased way of getting all the interactions," he says. "We have a flashlight illuminating 20% of the yard, but the other 80% is dark." In fact, no one yet knows how big the universe of interactions is, he says, "but everyone agrees that we are not even close to having mapped it".

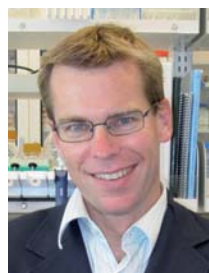
Nonetheless, more interactions have been

identified than can be individually investigated. For Ideker, the best approach is to think in terms of databases. "I have this big 'gamish' of interactions, how do I best query it?"

DATA COMBINATIONS

One strategy is combining diverse data sets around focused questions. For example, Ideker decided to conduct a Y2H screen that would pull out interactions involved in the mitogen-activated protein kinase (MAPK) signalling cascade — an important drug target that regulates processes such as cell growth, differentiation and survival. Ideker and his colleagues picked 150 proteins associated with the pathway and hunted for their interaction partners using Y2H assays. This revealed more than 2,000 interactions among about 1,500 proteins.

From these they selected a dozen or so proteins that had not previously been associated with the MAPK cascade and used RNA interference to knock down the expression of the identified interaction partners. In about one-third of the cases, RNA knockdown altered gene expression within the cascade, indicating that these interactions were functional. Follow-up studies provided the first experimental evidence that a protein called NHE-1 served as a MAPK scaffold¹⁰.



"It's the superposition of biophysical and functional data that is going to save the day."

Trey Ideker

By starting with the interactions and whittling them away with other data, the researchers can uncover new biology, says Ideker. "It's the superposition of biophysical and functional data that is really going to save the day here."

Researchers can also glean insight from how proteins interact physically. This year, Haiyuan Yu and his colleagues at Cornell University, Ithaca, New York, showed how combining data about protein-protein interactions and protein structure could suggest how certain mutations cause disease¹¹.

They combined several established data sets of protein-protein interactions, the physical structure of those interactions, and genetic measurements to show that when mutations do not prevent proteins from being expressed but still cause disease, they are more likely to occur in the interface between interacting proteins than elsewhere. "For the past decade, biologists have been using this mathematical definition. Every protein is a mathematical dot. But we know that protein structure is fundamentally important for function," says Yu.

Information about whether an interaction occurs in a specific cell type or under certain

conditions could go a long way to revealing its function, says Anne-Claude Gavin, who studies protein complexes at the European Molecular Biology Laboratory in Heidelberg, Germany. "Interactions have to be context-dependent; they have to start at one time and stop at another." But these studies are difficult and are rarely done. "This is a level of sophistication that we just don't understand," she says.

To understand a protein-protein interaction in context, researchers need to single them out for focused studies.

Sometimes, screening techniques can be adapted to follow particular interactions in depth. For example, complementation assays with fluorescent proteins or luciferase can be used to follow interacting proteins. Because different coloured fluorescent proteins are so similar, one protein can be tested for interactions with two or more proteins in the same cell. One protein is labelled with a fragment of yellow fluorescent protein, a second with a fragment of cyan fluorescent protein and another interrogated protein carries a fragment common to both fluorescent proteins. This can show which protein interactions are occurring and where in the cell they occur. Complementation assays with luciferase can also be used with multiple colours of proteins and have the advantage that the enzyme easily breaks apart and reforms, allowing researchers to study how interactions can be disrupted. Imaging techniques such as bioluminescent resonance energy transfer and fluorescent resonance energy transfer can be used in living cells. They use genetically tagged proteins that emit light when proteins come into contact with each other, and so are used in a variety of assays. Other assays label each of two proteins and then monitor whether they move together in cells.

Although slower and more expensive than large-scale screening efforts, one-at-a-time explorations of interactions are essential, says Uetz. "Eventually you want to drill down into the actual interactions." ■

Monya Baker is technology editor for *Nature* and *Nature Methods*.

1. Uetz, P. *et al. Nature* **403**, 623–627 (2000).
2. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. & Sakaki, Y. *Proc Natl Acad Sci USA* **98**, 4569–4574 (2001).
3. Venkatesan, K. *et al. Nature Methods* **6**, 83–90 (2009).
4. Fields, S. & Song, O. *Nature* **340**, 245–246 (1989).
5. Suter, B. & Wanker, E. E. (eds) *Two Hybrid Technologies: Methods and Protocols* (Humana, 2012).
6. Lievens, S. *et al. J. Proteome Res.* **8**, 877–886 (2009).
7. Barrios-Rodiles, M. *et al. Science* **307**, 1621–1625 (2005).
8. Braun, P. *et al. Nature Methods* **6**, 91–97 (2009).
9. Tarassov, K. *et al. Science* **320**, 1465–1470 (2008).
10. Bandyopadhyay, S. *et al. Nature Methods* **7**, 801–805 (2010).
11. Wang, X. *et al. Nature Biotechnol.* **30**, 159–164 (2012).
12. Jäger, S. *et al. Nature* **481**, 365–370 (2012).

CAREERS

COLUMN A tool for physicists to gauge risk and pick their projects wisely **p.279**

CAREERS BLOG The latest discussions and news on research jobs go.nature.com/ielkkf

NATUREJOBS For the latest career listings and advice www.naturejobs.com

3D STOCK



INNOVATION

The big idea of technology transfer

Working at the interface between science and business offers an opportunity to bring ideas to market.

BY CHARLOTTE SCHUBERT

As Angela Loihl worked her way through her graduate research project at the University of Iowa in Iowa City, she noticed the scope of her studies getting narrower. “I spent all my time learning more and more about less and less,” she says of her research in mice, which focused on a protein thought to have a role in stroke. “I questioned

how relevant that was to the human condition.”

Fourteen years after she earned her PhD, her career is far broader. As a technology manager at the Center for Commercialization at the University of Washington in Seattle, she covers a wide range of life-science fields — from microbiology to radiology, with an occasional foray into chemical engineering.

Loihl’s role is to take scientific ideas from academia and negotiate their transfer to

biotechnology start-ups and pharmaceutical companies, at which point the most promising leads may become new therapies. Loihl is routinely presented with intriguing, and sometimes thorny, questions. What is the commercial potential of a particular university technology? Which company will license the rights to which innovation? Are a researcher’s findings strong enough to launch a spin-off?

Loihl has helped to foster technology related to influenza treatments, tests to assess the risk of cardiovascular disease and a vaccine for the type 2 herpes simplex virus. “Every day I learn something new from the science and business perspective,” she says. “I love this job.”

Working in technology transfer demands teamwork and the ability to assess a huge range of scientific areas. Those who thrive have the right mixture of extroversion, scientific breadth and business sense. They are also able to juggle multiple projects at once. The hardest thing, says Loihl, is keeping everyone’s expectations realistic, such as ensuring a researcher does not overestimate the commercial value of his or her work, or that a company does not underestimate it. “If you do something well by one stakeholder, it usually means you are upsetting another,” says Loihl, adding that the variety of duties and challenges is what makes her job great.

WHERE THE JOBS ARE

Fourteen years ago, there were 26 employees at what was then called the office of technology transfer at the University of Washington. Now there are 54. The growth, which is not atypical for large academic institutions in the United States, is due in large part to the Bayh–Dole Act of 1980. The act changed the pace and manner of innovation in the United States, giving universities and not-for-profit organizations control over the intellectual-property rights of federally funded research done within the institution. The result was a huge increase in technology-transfer related opportunities, particularly in the 1990s as universities mined this new source of revenue.

The expanding offices have taken on functions beyond the usual roles of patenting and licensing technology, resulting in new types of work. At the Center for Commercialization, office head and entrepreneur Linden Rhoads leads a team of 15 technology managers (also known as licensing managers or licensing specialists), which are typical positions for people making the jump from science, as Loihl did. Rhoads also employs specialists who team ►

► up with the managers on various projects. “We have all kinds of people in support roles,” she says.

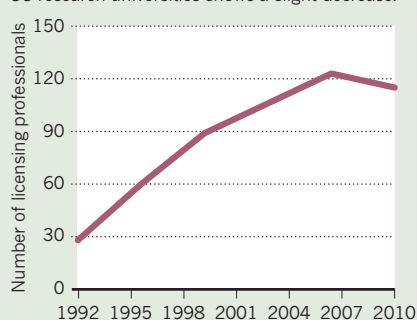
For instance, three industry-relations officers help the technology managers to understand industry needs and to forge collaborations between faculty members and companies. Most of these officers are scientists or engineers with extensive business experience. Four people, most of whom have business experience at small start-up companies, are in a ‘New Ventures’ group, which works with technology managers and researchers to start companies by inviting potential investors to the office, for example, and recruiting business people for an ‘entrepreneurs-in-residence’ programme. The New Ventures group also boasts a grant writer, who can pull in funding from sources such as the US Small Business Administration in Washington DC to foster a start-up.

Patent agents in Rhoads’ team, some of whom have law degrees, write the lengthy and technical patent applications and handle the continual back-and-forth between the patent office and researchers to keep the application moving forward — a task that technology managers often perform in smaller offices. Loihl advises job seekers to “scrutinize the office to see what kind of resources are available”.

Although Rhoads is still hiring staff, the big expansion in the number of US technology-transfer jobs has slowed, says Robin Rasor, director of licensing at the University of Michigan in Ann Arbor and former president of the Association of University Technology Managers in Deerfield, Illinois (see ‘Growth subsides’). The association’s annual survey of technology-transfer offices at US universities, hospitals and research institutions showed a slight drop in the number of

GROWTH SUBSIDES

Number of licensing professionals at eight major US research universities shows a slight decrease.



full-time licensing jobs from 1,041 to 1,018 between 2009 and 2010, at the 172 institutions that responded in both years. Job markets in Germany and the United Kingdom, which have long had laws similar to Bayh–Dole, are likewise slowing, says Anders Haugland, president of the Association of European Science and Technology Professionals, based in Leiden, the Netherlands. But ongoing changes in patent law and technology-transfer policies elsewhere have created job potential, notes Haugland, who also heads a technology-transfer office of seven institutions, including the University of Bergen and Haukeland University Hospital, in Norway.

Norway adopted policies similar to the Bayh–Dole Act in 2003. The countries universities now retain one-third of the proceeds from intellectual property arising from research conducted on their campuses. Previously, faculty members generally retained the full rights. Back then, Haugland’s technology-transfer office did not even exist. It now has 23 employees, some of whom have

come from as far as India, Croatia and Lithuania. “We are hiring people every year and we are growing,” says Haugland.

GETTING THE JOB

As Loihl became disillusioned with bench science, she began inquiring about internship opportunities at the then-small technology-transfer office at the University of Iowa, but they showed little interest. “They didn’t know how they would use me,” she says.

But times have changed. Technology-transfer offices now offer various paid and unpaid internship programmes. The University of Washington has a typical technology-licensing intern programme. Interns prepare summaries about the university researcher’s innovations to facilitate licensing negotiations, and help to evaluate whether a new technology can be patented and whether there is a potential market for it. Often an internship can lead to a longer, salaried apprenticeship, says Rhoads. A technology-transfer role may also lead to a position in industry (see ‘Jumping to industry’).

An internship is the best way to get a foot in the door, especially for those making the transition from science, says Rasor. She and Rhoads have both hired interns who came straight from postdoctoral fellowships or graduate programmes as junior technology managers.

Courses in business, marketing or law can help to land a job or internship. PraxisUnico, a not-for-profit organization in Cambridge, UK, offers a three-day ‘fundamentals of technology transfer’ course, which is popular for those new to the sector, says Alison Campbell, a consultant and chair of PraxisUnico’s training committee. Aspiring technology managers can learn the basics of patenting, trademarks and licensing, as well as marketing and negotiation tactics.

Although not all of Rhoads’ team have a masters in business, she says that the degree can help with tasks such as assessing the market potential for a technology — for example, by researching the size of the market, cost of goods or pricing models. Rhoads herself has a law degree, which helps in understanding the complicated language of contracts. But in the end, it may be business experience that carries the most weight. “If you have someone who has any kind of experience in industry, that is a big plus,” says Rasor. Experience in a business office is preferable, but even work as a researcher at biotechnology company can help, she says.

There is no typical route into a technology-transfer role. And ultimately, less-tangible abilities, such as communication and people skills, could be what leads to success in the field. “People need to be good at having consultations with scientists, turning over rocks and introducing researchers from industry to academic researchers,” says Rhoads. Those who thrive in the job, she says, are “natural connectors”. ■

Charlotte Schubert is a freelance journalist based in Seattle, Washington.

JUMPING TO INDUSTRY

Making the move to the private sector

Working in a technology-transfer office does not mean staying in academia. Some use their technology-transfer experience to make the leap to the private sector.

“We are at the other side of the deals,” says Polly Murphy, a vice-president at drug firm Pfizer in San Diego, California, who leads 30 people in the business development team for the company’s research and development branch. The team negotiates transactions with other companies, not-for-profit organizations, universities and other institutions. Most of the team came from the technology-transfer world, which, says Murphy, is the best place to learn how to draft licensing and collaboration agreements. Working in technology transfer can also lead to a job in

a contracts and outsourcing office, which manages agreements for clinical trials and other projects. But Murphy cautions that shrinkage in the industry has made such jobs scarce. “In my department nobody ever leaves,” she says.

Other technology-transfer employees head to biotechnology firms or use their scientific knowledge and understanding of the market to work in a venture-capital firm. But biotechnology start-ups come and go, along with their licensing offices. Those who are risk averse should be wary of the jump to industry, says Anders Haugland, head of the Association of European Science and Technology Professionals in Leiden, the Netherlands. And universities and research institutions often have better benefits. **C.S.**

COLUMN

Rating research risk

Too many young physicists embark on projects without knowing the risks. There is a better way, argues **Abraham Loeb**.

In physics, the value of a theory is measured by how well it agrees with experimental data. But how should the physics community gauge the value of an emerging theory that cannot yet be tested experimentally? With no reality check, a less than rigorous hypothesis such as string theory may linger on, even though physicists have been unable to work out its actual value in describing nature.

This sort of uncertainty has implications not only for the gathering of knowledge for the scientific enterprise, but also for fledgling physicists. The investment of research time in strong intellectual assets is crucial for graduate students who want to establish their careers on a good foundation. But not all young researchers are aware of the history that accompanies every research area. They often have to rely on word of mouth from their PhD adviser or colleagues for this information.

What if the physicists could call on a ratings agency, not unlike a lender would do before deciding whether to offer credit? I am advocating the creation of a website that is operated by graduate students and that will use various measures of publicly available data (such as the number of newly funded experiments, research grants, publications and faculty jobs) to gauge the future returns of various research frontiers.

THEORY BUBBLES

The study of the cosmic microwave background provides an example of how theory and data can generate opportunities for young scientists. As soon as NASA's Cosmic Background Explorer satellite reported conclusive evidence for the cosmic microwave background temperature fluctuations across the sky in 1992, the subsequent experimental work generated many opportunities for young theorists and observers who joined this field. By contrast, a hypothesis such as string theory, which attempts to unify quantum mechanics with Albert Einstein's general theory of relativity, has so far not been tested critically by experimental data, even over a time span equivalent to a physicist's career.

Senior scientists might seem the people best suited to rate the promise of research frontiers. But too many of these physicists are already invested in evaluating the promise of these speculative theories, implying that they could have a conflict of interest or be wishful thinkers. Having these senior scientists rate future promise would be akin to the 'AAA' rating that financial



agencies gave to the very debt securities from which they benefited. This unseemly situation contributed to the last recession, and a long-lived bias of this type in the physics world could lead to similarly devastating consequences — such as an extended period of intellectual stagnation and a community of talented physicists investing time in research ventures unlikely to elucidate our understanding of nature — a theory 'bubble', to borrow from the financial world.

Of course, graduate students are busy. But they could serve a limited term of service for maintaining the site and be government-funded. For example, students supported by US National Science Foundation fellowships maintain astrobits.com, a website that summarizes new astrophysics papers.

CREDIT RATING

The physics 'credit-rating' website would use evaluation metrics to factor in, with the correct weighting, all the ingredients that would ultimately make scientific research successful.

For physics, this might include the existence of an underlying self-contained theory from first principles, the potential for experimental tests of this theory and a track record of related research programmes. Clearly, factors such as intellectual excitement cannot be quantified, but as long as funding agencies are supporting projects and the information provided is accurate, the data about the growth of a field should echo this 'excitement' factor.

The evaluation metric would have to be predetermined and supported by numbers that are based on archival data gathered through automated searches for keywords in electronic data

archives (see arxiv.org or nsf.org). Aside from automated searches, practitioners from fields that are being evaluated could submit supplementary data that would be incorporated into the analysis.

The entire data set would include the level of funding allocated to experiments and research grants, the status of the underlying theory and the number of publications and faculty jobs within the particular field of research. The simplest model relates the change in these parameters to a linear combination of their values. For example, the publication rate is expected to relate to a linear combination of the number of experiments, faculty jobs and available research funds. With the right mix of time spent on theory, experimental work and grant support, a research frontier would show exponential growth in this linear model. The next step would be to calibrate this model using historical data about the growth of successful research frontiers.

The website could be helpful to institutions and governments, not just to individual scientists. A balanced assessment of the level of risk and potential benefits from emerging research frontiers can increase the efficiency of the workforce, leading to stronger growth. And it could help funding agencies to optimize their allocation of money to promote progress in research. In fact, it would be in the interests of funding agencies to support the website and help the students to take part (for example, through special grants or fellowships).

The website might also convince senior researchers to shift their focus to new research areas, perhaps as a result of the influence that the rating procedure may have on funding agencies. But maintaining balance and ensuring diversity among subfields, taking some risks and avoiding funnelling resources into a small number of successful but conservative programmes are important considerations for funding agencies (A. Loeb *Nature* **467**, 358; 2010).

Nearly every worthwhile endeavour involves some risk. But mitigating that risk, and helping young scientists to make informed decisions about the field in which they should invest their time and intellect, would yield a more efficient scientific enterprise. ■

Abraham Loeb is chair of the astronomy department and director of the Institute for Theory and Computation at Harvard University in Cambridge, Massachusetts.

NOR CUSTOM STALE

An antic disposition.

BY ANATOLY BELILOVSKY

Old people move slowly. A knee may work just fine today and buckle tomorrow. A tiny turn can change a functioning joint into a monolithic block of agony.

Old people need to remember things like that.

"You look great," said Bob, looking down and chewing at his dentures.

Bob had been a good striker, back in the day, but he never learned to fool goalies. He always looked where he wanted the ball to go. So in the end, I went pro, and Bob went into insurance.

"Really good," Bob added. His feelings littered his face the way his lunch decorated his tie: a forced tight, guilty smile with pity in the tilt of his head, and a tiny trace of *Schadenfreude* in the crow's feet around his eyes.

I knew how I looked. I looked terrible: slumped shoulders, shuffling gait, shabby jacket, stained pants.

"Thanks," I whispered. "Look at us. Sixtieth reunion. Who'd'a think we'd live this long, eh, roomie?"

Bob looked up, his smile genuine now. "Athletes don't age as fast," he said. "I still play nine holes, every week. Like clockwork. Keeps me young."

I nodded. "It's working," I whispered. "You haven't changed a bit."

He barked out a single laugh. "Right. Not a bit," he said. Leaning back, he turned his head slowly through a short arc, sweeping his gaze over the far side of the quad. "Princeton sure has changed," he continued. "There's dorms where soccer fields used to be. Where you and I played. Mixed dorms! Not just co-ed. Mixed. Ain't that something!" He slapped his knee, winced slightly.

I nodded again.

"Trouble with your voice?" he asked.

"Sort of," I whispered.

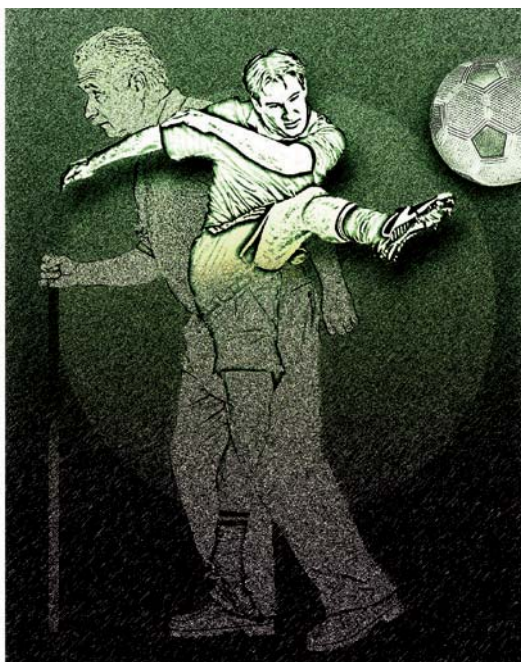
He sighed. "Joanie had a stroke, she can't talk at all," he said. "And Todd died last year of throat cancer, he had a tracheostomy. Had to plug the hole in his neck when he wanted to say anything."

I sighed, too, and lowered my eyes. Joanie and I lived together, sophomore year. I'd never fool her for a minute.

"I worked as long as I could, only retired after I got my bypass," he said. His face took on that look again: guilt, pity and a dash of gloat. Same look he had when I'd told him

Joanie had left me. Minus the dentures and the hand tremor. "Joanie gets round-the-clock nursing. Government pays for everything," he added. "Wouldn't wanna get old in a country where you pay for medicine."

Joanie moved in with him a week later. Maybe that was why I went overseas, so I'd never have to see that look again. In either of their faces.



Bob reached for my shoulder, his fingers trembling as if rolling an invisible cigarette. I fought the urge to move away.

"I heard health care is expensive where you live," he said.

I nodded, with far more force than I expected.

Bob looked up sharply.

I winced and rubbed my neck a fraction of a second later. Bob leaned closer to me. His shiny dentures clashed with his cracked, spit-flecked lips; his eyes, once brown, were now ochre iris on yellow sclera.

"You could get your citizenship back," he said, barely moving his lips. "My grandson is a damn good immigration lawyer."

I shrugged.

"Think about it," he said.

ON NATURE.COM

Follow Futures on

Facebook at:

go.nature.com/mtoodm

"I will," I said. "But now I have to go. Don't want to miss my flight."

"Stay with me," he said quickly. "We've

guest rooms at the home. We can go see Joanie tomorrow."

I counted to five and held my breath: an old voice actor's trick.

"No," I said.

The word came out as I intended: with longing. With reluctance. With regret.

Bob shook my hand and hobbled away as quickly as he could.

I had to sprint across the terminal to reach my gate on time. People stared. My wrinkles itched; I detoured to the lavatory to peel them off. The TranSec agent looked at me with a suspicious squint.

I called the Farm from the plane while it waited for clearance. Gulnara answered.

"Well, hello, stranger," she purred. "Do we have a date?"

"Sure," I said. "When can you fit me in?"

"You want a quickie," she said, "or the whole jalapeño?"

"Enchilada," I said. Gulnara's English was perfect. It was American she had trouble with. "I want the whole enchilada. It's been a while."

"So I see." She paused; I heard keys tapping. "I have an opening Wednesday. Is this good?"

"Sure," I said. "Training camp doesn't start for another fortnight."

I heard tapping again. "Excellent. Full rejuvenation, a five-day course starting Wednesday, shall I debit your fee now?"

"Go ahead," I said. "Did it go through?"

"With your credit rating?" she said. "Of course it went through." She paused. "I'm so glad you haven't retired. Watching you play — it never gets old. It's like, you are not just playing soccer, but also poker and chess at the same time. Does this make sense?"

"Sweetheart," I said, "I can't afford to retire."

Her answer drowned in the turbine spin-up. I disconnected my phone, leaned back, turned on the viewscreen. The plane made a climbing turn above central Jersey before heading over the Atlantic. Somewhere below, shabby, weed-choked Princeton sweltered in the heat, and Bob shambled with a cane to the train that would take him to his nursing home.

Pity about poor Bob. ■

Anatoly Belilovsky learned English from Star Trek reruns and is now a paediatrician in New York.

The clonal and mutational evolution spectrum of primary triple-negative breast cancers

Sohrab P. Shah^{1,2*}, Andrew Roth^{1,2*}, Rodrigo Goya^{3*}, Arusha Oloumi^{1,2*}, Gavin Ha^{1,2*}, Yongjun Zhao^{3*}, Gulisa Turashvili^{1,2*}, Jiarui Ding^{1,2*}, Kane Tse^{3*}, Gholamreza Haffari^{1,2*}, Ali Bashashati^{1,2*}, Leah M. Prentice^{1,2}, Jaswinder Khattri^{1,2}, Angela Burleigh^{1,2}, Damian Yap^{1,2}, Virginie Bernard⁴, Andrew McPherson^{1,2}, Karey Shumansky^{1,2}, Anamaria Crisan^{1,2}, Ryan Giuliani^{1,2}, Alireza Heravi-Moussavi^{1,2}, Jamie Rosner^{1,2}, Daniel Lai^{1,2}, Inanc Birol³, Richard Varhol³, Angela Tam³, Noreen Dhalla³, Thomas Zeng³, Kevin Ma³, Simon K. Chan³, Malachi Griffith³, Annie Moradian³, S.-W. Grace Cheng³, Gregg B. Morin^{3,5}, Peter Watson^{1,6}, Karen Gelmon⁶, Stephen Chia⁶, Suet-Feung Chin^{7,8}, Christina Curtis^{7,8,9}, Oscar M. Rueda^{7,8}, Paul D. Pharoah⁷, Sambasivarao Damaraju¹⁰, John Mackey¹⁰, Kelly Hoon¹¹, Timothy Harkins¹¹, Vasisht Tadigotla¹¹, Mahvash Sigaroudinia¹², Philippe Gascard¹², Thea Tlsty¹², Joseph F. Costello¹³, Irmtraud M. Meyer^{5,14,15}, Connie J. Eaves¹⁶, Wyeth W. Wasserman^{4,5}, Steven Jones^{3,5,17}, David Huntsman^{1,2,18}, Martin Hirst^{3,15,19}, Carlos Caldas^{7,8,20,21}, Marco A. Marra^{3,5} & Samuel Aparicio^{1,2}

Primary triple-negative breast cancers (TNBCs), a tumour type defined by lack of oestrogen receptor, progesterone receptor and *ERBB2* gene amplification, represent approximately 16% of all breast cancers¹. Here we show in 104 TNBC cases that at the time of diagnosis these cancers exhibit a wide and continuous spectrum of genomic evolution, with some having only a handful of coding somatic aberrations in a few pathways, whereas others contain hundreds of coding somatic mutations. High-throughput RNA sequencing (RNA-seq) revealed that only approximately 36% of mutations are expressed. Using deep re-sequencing measurements of allelic abundance for 2,414 somatic mutations, we determine for the first time—to our knowledge—in an epithelial tumour subtype, the relative abundance of clonal frequencies among cases representative of the population. We show that TNBCs vary widely in their clonal frequencies at the time of diagnosis, with the basal subtype of TNBC^{2,3} showing more variation than non-basal TNBC. Although *p53* (also known as *TP53*), *PIK3CA* and *PTEN* somatic mutations seem to be clonally dominant compared to other genes, in some tumours their clonal frequencies are incompatible with founder status. Mutations in cytoskeletal, cell shape and motility proteins occurred at lower clonal frequencies, suggesting that they occurred later during tumour progression. Taken together, our results show that understanding the biology and therapeutic responses of patients with TNBC will require the determination of individual tumour clonal genotypes.

To understand the patterns of somatic mutation in TNBC, we enumerated genome aberrations at all scales from 104 cases of primary TNBC (Affymetrix SNP6.0, 104 cases; RNA-seq, 80 cases; genome/exome sequencing, 65 cases: 54 exomes, 15 genomes with 4 overlapping) (Supplementary Table 1 and Supplementary Fig. 1), annotated with clinical information (Supplementary Table 2). We revalidated 2,414 somatic single nucleotide variants^{4,5} (SNVs) (Supplementary Table 3) with targeted deep sequencing to a median of 20,000×

coverage, including 43 non-coding splice site dinucleotide mutations (Supplementary Table 4) and 104 genes with 107 indels (Supplementary Table 5 and Supplementary Methods). Notably, the distribution of somatic mutation abundance varies in a continuous distribution among tumours (Fig. 1a) and seems to be unrelated to the proportion of the genome altered by copy number alterations (CNAs) (Fig. 1b) or tumour cellularity (Supplementary Fig. 2b). Although this distribution could be partially explained by a false-negative rate in mutation discovery, others have noted similar distributions in epithelial cancers⁶, suggesting that the total mutation content of individual tumours may be shaped by biological processes or differential exposure to mutagenic influences in the population.

The overall pattern (Supplementary Fig. 3a, b) of CNA abundance appears similar (Supplementary Fig. 4) to that seen in a larger, independent series of ~2,000 SNP6.0 profiled breast tumours⁷. Among the most frequently observed CNA events (Supplementary Table 6) are the tumour suppressor and oncogenes *PARK2* (6%), *RB1* (5%), *PTEN* (3%) and *EGFR* (5%). Here we report intragenic deletions (Supplementary Fig. 5) in the *PARK2* tumour suppressor^{8,9}, specifically linking *PARK2* with TNBC for the first time. Consistent with previous reports in breast cancer¹⁰, we did not observe frequent recurrent structural rearrangements (Supplementary Fig. 3d and Supplementary Table 7), although we revalidated many individual fusion events involving known oncogenes or tumour suppressors (for example, *KRAS*, *RB1*, *IDH1*, *ETV6*) (Supplementary Tables 8–10).

A comparison of RNA-seq data with genomes/exomes data revealed that only 36% of validated somatic SNVs were observed in the transcriptome sequence (Supplementary Table 3 and Supplementary Fig. 2b). In a recent lymphoma study, similar proportions were observed (137 of 329 somatic mutations expressed in RNA-seq)¹¹. As expected, the proportion of low-abundance somatic SNVs observed in RNA is reflected in the distribution of wild-type, heterozygous and homozygous expressed mutations (Supplementary Fig. 2b), consistent with the notion that

¹Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada. ²Molecular Oncology, British Columbia Cancer Research Centre, Vancouver, British Columbia V5Z 1L3, Canada. ³Canada's Michael Smith Genome Sciences Centre, Vancouver, British Columbia V5Z 1L3, Canada. ⁴Centre for Molecular Medicine and Therapeutics, 950 West 28th Avenue, Vancouver, British Columbia V5Z 4H4, Canada. ⁵Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada. ⁶British Columbia Cancer Agency, 600 West 10th Avenue, Vancouver, British Columbia V5Z 4E6, Canada. ⁷Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. ⁸Department of Oncology, University of Cambridge, Hills Road, Cambridge CB2 2XZ, UK. ⁹Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA. ¹⁰Departments of Oncology and Laboratory Medicine and Pathology, University of Alberta, 11560 University Avenue, Cross Cancer Institute, Edmonton, Alberta T6G 1Z2, Canada. ¹¹Life Technologies, 101 Lincoln Centre Dr., Foster City, California 94404, USA. ¹²Department of Pathology and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94143, USA. ¹³Brain Tumor Research Center, Department of Neurosurgery, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California 94143, USA. ¹⁴Department of Computer Science, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada. ¹⁵Centre for High-Throughput Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada. ¹⁶Terry Fox Laboratory, BC Cancer Agency, 675 W 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada. ¹⁷Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, British Columbia V5A1S6, Canada. ¹⁸Centre for Translational and Applied Genomics, BC Cancer Agency, 600 West 10th Ave, Vancouver, British Columbia V5Z 4E6, Canada. ¹⁹Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada. ²⁰Cambridge Breast Unit, Addenbrookes Hospital, Cambridge University Hospital NHS Foundation Trust and NIHR Cambridge Biomedical Research Centre, Cambridge CB2 2QQ, UK. ²¹Cambridge Experimental Cancer Medicine Centre (ECMC), Cambridge CB2 0RE, UK.

*These authors contributed equally to this work.

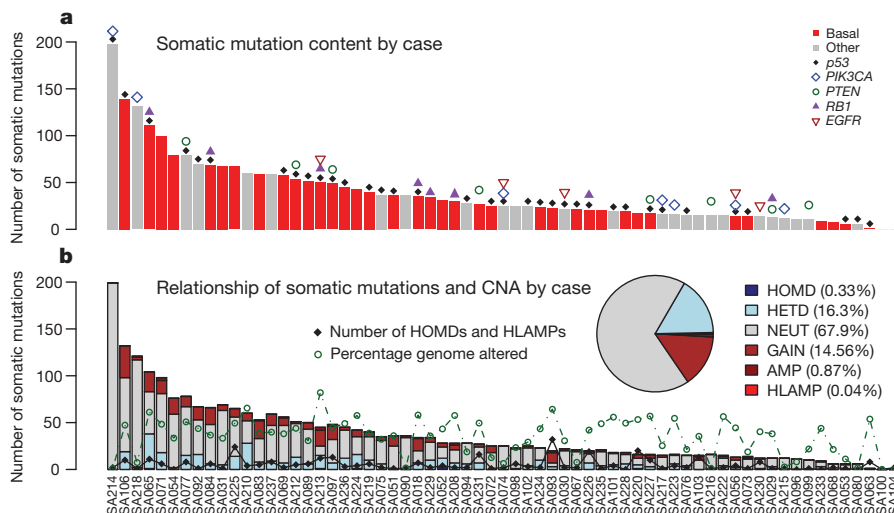


Figure 1 | Distribution of number of validated somatic mutations by case over 65 cases. **a**, Mutation frequency (basal, red; other, grey). Patients harbouring known driver gene mutations are indicated. **b**, Case-specific and overall (inset) distributions of mutations in CNA classes. AMP, amplification; GAIN, single copy gain; HETD, hemizygous deletion; HLAMP, high-level amplification; HOMD, homozygous deletion; NEUT, no copy number change. The number of (HOMD, HLAMP) CNAs (black diamonds) and percentage genome altered (green circles) are indicated.

low-abundance alleles may represent rarer clones in the primary tumour. We found 43 splice junction mutations with evidence for an impact on splicing patterns (Supplementary Table 4), encompassing several known tumour suppressors (*p53*, *PIK3R1*; Supplementary Fig. 6) as well as many genes not yet implicated in carcinogenesis. Analysis of 72 somatic mutations in the non-coding space of experimentally determined human regulatory regions¹² showed (Supplementary Table 11) a significant overrepresentation (31.9% versus expected 2.5%, Fisher exact test $P = 2 \times 10^{-19}$) of mutations within retinoblastoma-associated protein (RB)-binding sites. Six mutations were predicted to be damaging to RB binding (Supplementary Methods and Supplementary Fig. 7). This is consistent with observations of frequent functional disruption of the RB-regulated cell cycle network¹³ in TNBC.

We next searched for mutation enrichment patterns in three ways: by single gene mutation frequency over multiple cases; by the mutation frequency over multiple members of a gene family; and by correlating mutation status with expression networks. First, similar to other studies^{14,15}, *p53* is the most frequently mutated gene (Supplementary Table 12) with 62% of basal TNBC (determined by gene expression classification with PAM50 (ref. 16) analysis on RNA-seq expression profiles) and 43% of non-basal TNBC cases harbouring a validated somatic mutation. We also observed frequent mutations in *PIK3CA* at 10.2% (7/65), *USH2A* (Usher syndrome gene, implicated in actin cytoskeletal functions) at 9.2% (6/65), *MYO3A* at 9.2%, *PTEN* and *RB1* at 7.7% (5/65) and a further eight genes (including *ATR*, *UBR5* (also known as *EDD1*), *COL6A3*) at 6.2% (4/65) of cases in the cohort

(Fig. 2a). Considering background mutation rates¹⁷, *p53*, *PIK3CA*, *RB1*, *PTEN*, *MYO3A* and *GHI* showed evidence of single gene selection ($q < 0.1$) (Supplementary Table 13). Additional recurrent mutations of note occurred in the synuclein genes (*SYNE1* and *SYNE2*, 9.2% 6/65, recently implicated in squamous head and neck cancers^{18,19}), *BRCA2* (three cases), and several other well known oncogenes (*BRAF*, *NRAS*, *ERBB2* and *ERBB3*) with mutations in two cases each. Approximately 20% of cases contained examples of potentially 'clinically actionable' somatic aberrations, including *BRAF* V600E, high-level *EGFR* amplifications and *ERBB2* and *ERBB3* mutations.

In the second approach we searched for statistically overrepresented gene families and protein functions using the Reactome functional protein interaction database²⁰ (Supplementary Methods). This analysis quantifies gene family involvement through sparse mutation patterns in functionally connected genes, which would be statistically underrepresented by single gene recurrent mutation analysis. The overrepresented pathways (false discovery rate (FDR) < 0.001) included *p53*-related pathways along with chromatin remodelling, PIK3 signalling, ERBB signalling, integrin signalling and focal adhesion, WNT/cadherin signalling, growth hormone and nuclear receptor co-activators, and ATM/RB-related pathways (Fig. 3a and Supplementary Table 14). We note that the candidate 'driver' *MYO3A*, a cytoskeleton motor protein involved in cell shape and motility, relates to several pathways upstream and downstream of integrin signalling. The mutated genes include extracellular matrix (ECM) interactions (laminins, collagens), ECM receptors (integrins), several proteins

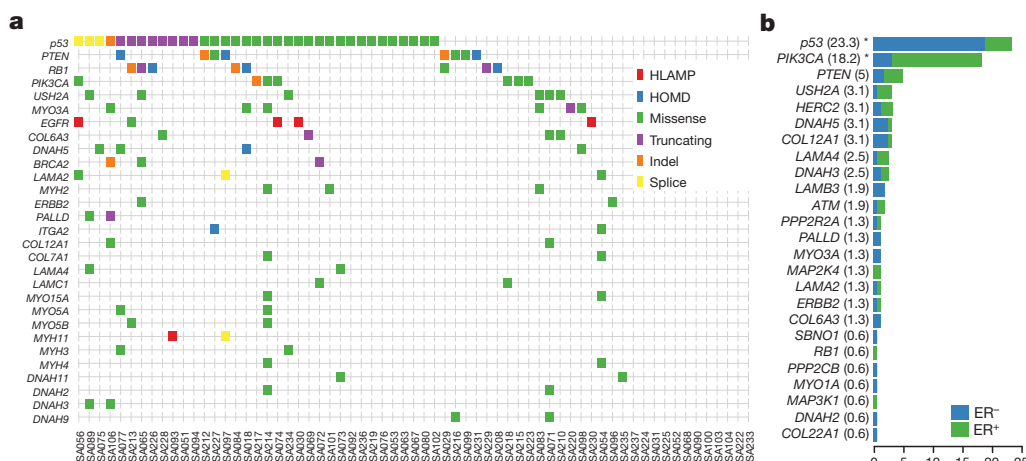


Figure 2 | Population patterns of co-occurrence and mutual exclusion of genomic aberrations in TNBC. **a**, Case-specific mutations in known driver genes, plus genes from integrin signalling and ECM-related proteins (laminins, collagens, integrins, myosins and dynein) derived from all aberration types: high-level amplifications (HLAMP), homozygous deletions (HOMD), missense, truncating, splice site and indel somatic mutations are depicted in genes with at least two aberrations in the population. **b**, Distribution of somatic mutations in 25 genes across all exons of 159 additional breast cancers (relative proportion of ER⁺ cases in green, and ER⁻ in blue), shown as a percentage of cases (in parentheses) with one or more mutations. * $P < 0.05$.

regulating actin cytoskeleton dynamics (usherin, palladin, multiple myosins) and microtubule motor proteins (kinesins) (Fig. 2a). All of these contribute to cellular processes that have been functionally implicated in cancer progression; however, a signature of somatic mutation associated with these proteins has not been previously noted in TNBC.

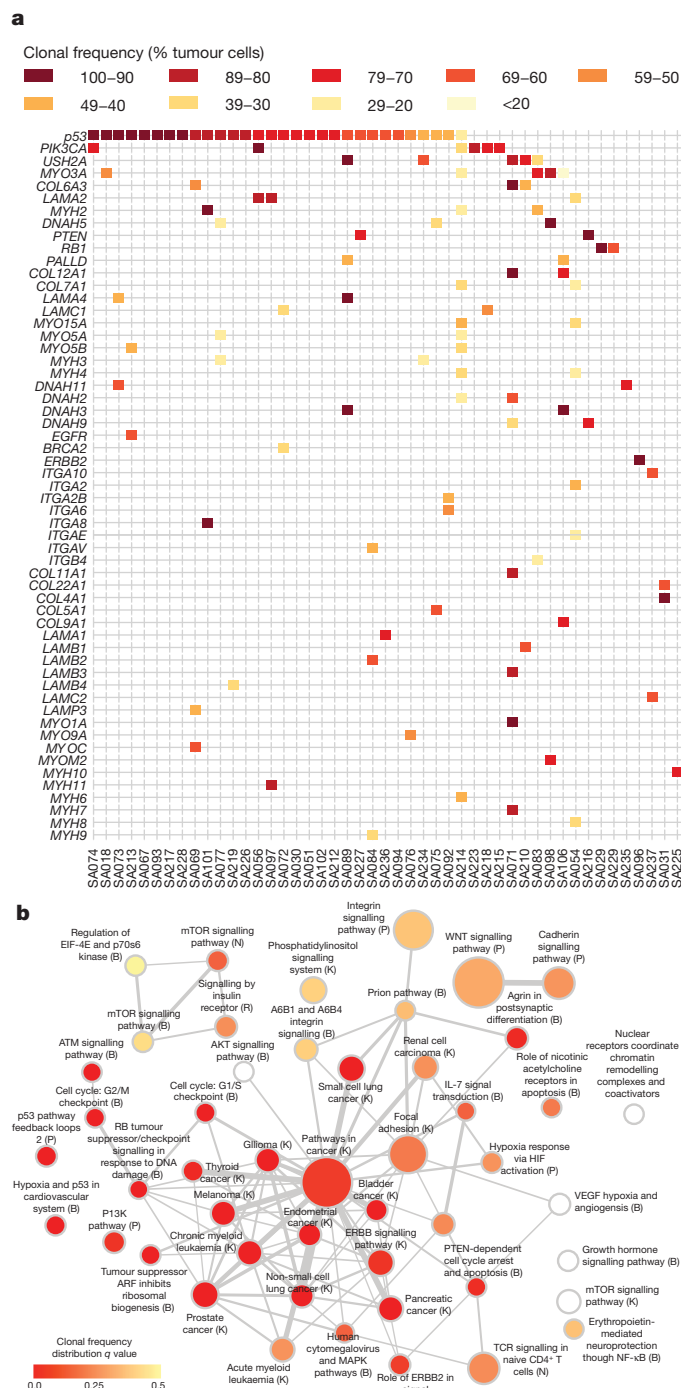


Figure 3 | Network analysis of 254 recurrently mutated genes by somatic point mutations and indels. a, Case-specific mutations shaded according to clonal frequencies in known driver genes, plus genes from integrin signalling and ECM-related proteins (laminins, collagens, integrins, myosins and dyneins). **b**, Significantly overrepresented pathways (FDR < 0.001) from recurrently mutated genes (see Supplementary Methods). Node shading encodes the adjusted *P* value (*q* value) of the comparison of the distribution of clonal frequencies of mutations in a given pathway to the overall distribution of clonal frequencies. A spectrum of higher (red) and lower (yellow) clonal frequencies is evident. Letters in parentheses indicate database sources.

To confirm the mutational spectrum in the general breast cancer population we re-sequenced all exons of 29 genes in an additional 159 breast cancers (82 oestrogen receptor (ER)⁺ and 77 ER⁻, tumour and matched normal) (Fig. 2b), and confirmed that many of the genes found in the discovery cohort were recurrently mutated in an additional population. Whether this pattern of mutation represents the occurrence of disease-modifying mutations, or possibly selection from other processes (for example, transcription-related hypermutation) is unknown. Interestingly, the enrichment of cytoskeletal functions in the somatic aberration landscape is also evident from the copy number and alternative splicing landscapes (Supplementary Fig. 8).

Third, we integrated both the CNA and mutation data with expression data to reveal genomic events associated with extreme changes in the transcription of interacting genes²⁰ (Table 1), using a bipartite graph-based method (driverNet; Supplementary Methods). The somatic aberrations showing statistically significant association with extreme expression in this analysis (*P* < 0.05) (Table 1 and Supplementary Table 15) implicate well known oncogenes and tumour suppressors (*TP53*, *PIK3CA*, *NRAS*, *EGFR*, *RB1*, *ATM*) and suggest several new genes of interest, including *PRPS2* (a nucleotide biosynthesis enzyme, rank 7), harbouring homozygous deletions in three cases, *NRC31* (a glucocorticoid receptor, rank 10) with SNVs in three cases, four PKC-related genes, *PRKCZ*, *PRKCQ*, *PRKG1* and *PRKCE*. The gene networks show a partial overlap with driverNet applied to the TCGA ovarian high-grade serous data²¹ (Supplementary Table 16).

Having identified candidate driver genes and significantly over-represented pathways, we asked how these are distributed among individual tumours by clustering a pathway–patient–mutation matrix (Supplementary Fig. 9). The abundance of implicated pathways can be

Table 1 | Analysis of the top somatically aberrated genes influencing expression

| Rank | Gene | gband | SNV or indel | HLAMP | HOMD | Events | <i>P</i> value |
|------|----------------|----------|--------------|-------|------|--------|--------------------|
| 1 | <i>TP53</i> | 17p13.1 | 35 | 0 | 0 | 2242 | 0 |
| 2 | <i>PIK3CA</i> | 3q26.32 | 7 | 0 | 0 | 441 | 1×10^{-4} |
| 3 | <i>NRAS</i> | 1p13.2 | 2 | 0 | 0 | 271 | 4×10^{-4} |
| 4 | <i>EGFR</i> | 7p11.2 | 1 | 5 | 0 | 220 | 4×10^{-4} |
| 5 | <i>RB1</i> | 13q14.2 | 5 | 0 | 5 | 184 | 5×10^{-4} |
| 6 | <i>PGM2</i> | 4p14 | 1 | 0 | 1 | 172 | 5×10^{-4} |
| 7 | <i>PRPS2</i> | 23p22.2 | 0 | 0 | 3 | 171 | 5×10^{-4} |
| 8 | <i>PTEN</i> | 10q23.31 | 5 | 0 | 3 | 150 | 5×10^{-4} |
| 9 | <i>PRKCE</i> | 2p21 | 0 | 0 | 1 | 136 | 7×10^{-4} |
| 10 | <i>NRC31</i> | 5q31.3 | 3 | 0 | 0 | 130 | 7×10^{-4} |
| 11 | <i>CREBBP</i> | 16p13.3 | 1 | 0 | 1 | 119 | 8×10^{-4} |
| 12 | <i>CS</i> | 12q13.2 | 1 | 0 | 0 | 108 | 0.0011 |
| 13 | <i>MAN2A2</i> | 15q26.1 | 2 | 0 | 1 | 104 | 0.0012 |
| 14 | <i>HMGCS2</i> | 1p12 | 1 | 2 | 0 | 100 | 0.0013 |
| 15 | <i>HEXA</i> | 15q24.1 | 2 | 1 | 0 | 97 | 0.0013 |
| 16 | <i>ADCY9</i> | 16p13.3 | 2 | 1 | 0 | 91 | 0.0017 |
| 17 | <i>OR4N4</i> | 15q11.2 | 0 | 0 | 5 | 90 | 0.0017 |
| 18 | <i>LCLAT1</i> | 2p23.1 | 0 | 0 | 1 | 85 | 0.002 |
| 19 | <i>DGKI</i> | 7q33 | 2 | 0 | 0 | 82 | 0.0022 |
| 20 | <i>CYP2A6</i> | 19q13.2 | 1 | 0 | 0 | 80 | 0.0024 |
| 21 | <i>JAK1</i> | 1p31.3 | 1 | 0 | 0 | 78 | 0.0026 |
| 22 | <i>POLR1A</i> | 2p11.2 | 2 | 0 | 0 | 78 | 0.0026 |
| 23 | <i>PLD1</i> | 3q26.31 | 1 | 0 | 0 | 69 | 0.0038 |
| 24 | <i>IDH3B</i> | 20p13 | 1 | 0 | 1 | 68 | 0.004 |
| 25 | <i>PAPSS2</i> | 10q23.2 | 0 | 0 | 3 | 67 | 0.0041 |
| 26 | <i>PRKX</i> | 23p22.33 | 0 | 0 | 2 | 65 | 0.0046 |
| 27 | <i>TPH2</i> | 12q21.1 | 1 | 0 | 0 | 65 | 0.0046 |
| 28 | <i>UGT2B17</i> | 4q13.2 | 0 | 0 | 1 | 63 | 0.0053 |
| 29 | <i>RRM2</i> | 2p25.1 | 1 | 0 | 0 | 57 | 0.0072 |
| 30 | <i>ATM</i> | 11q22.3 | 1 | 0 | 0 | 55 | 0.0084 |
| 31 | <i>CLCA1</i> | 1p22.3 | 2 | 0 | 0 | 54 | 0.009 |
| 32 | <i>PRKCZ</i> | 1p36.33 | 1 | 0 | 0 | 53 | 0.0095 |

Rank, derived by the driverNet algorithm (see Supplementary Methods); gene, somatically aberrated gene; gband, chromosomal band containing gene; SNV or indel, the number of cases harbouring an SNV or indel in the gene; HLAMP, the number of cases harbouring a predicted high-level amplification; HOMD, the number of cases harbouring a predicted homozygous deletion; events, number of gene expression outliers (see Supplementary Methods) coincident with a genomic aberration and where the outlying gene is connected to the aberrated gene; *P* value, statistical significance based on a randomly generated background distribution (Supplementary Methods).

seen to be only partially related to the total number of mutations in a case, groups 1 and 2 having on average fewer mutations per case. The frequent involvement of pathways with *p53*, *PTEN* and *PIK3CA* as members, is noted (Supplementary Fig. 9); however, the case groupings also vary by the progressive inclusion of additional pathways (for example, WNT signalling, integrin signalling, ERBB signalling, hypoxia and PI3K). More than two thirds of cases contained one or more mutations in the actin/cytoskeletal functions group of genes (Supplementary Fig. 9). Some 12% of cases did not contain somatic aberrations in any of the frequent drivers or cytoskeletal genes (Supplementary Table 12). This suggests that primary TNBCs are mutationally heterogeneous from the outset, with some patients' tumours having a small number of implicated pathways and few mutations, whereas other patients present with tumours containing extensive mutation burdens and multiple pathway involvement.

Motivated by the observation that early primary TNBCs show a wide variation of mutation content, we asked whether the clonal composition of these primary cancers is similarly varied. We and others have shown^{22,23} how deep-frequency measurements of allelic abundance can be used to study tumour clonal evolution. Clonal mutation frequency, a compound measure of clonal complexity, (Fig. 4a) can be estimated from allele abundance, once the influence of copy number states, regional loss of heterozygosity (LOH state) and tumour cellularity have been considered (although we note that approximately 68% of SNVs in this study are in diploid, neutral regions). To extend allelic abundance measurements to estimation of clonal frequencies, we implemented a Dirichlet process clustering model (pyclone; Supplementary Methods and Supplementary Fig. 10) that simultaneously estimates the genotype and clonal frequency given a list of deeply sequenced mutations and their local copy number and heterozygosity contexts.

Using the set of deeply sequenced (median 20,000×), validated SNVs, our analysis revealed (Fig. 4b) that groups of mutations within individual cases have different clonal frequencies, indicative of distinct clonal genotypes. Remarkably, the tumours exhibit a wide spectrum of modes over clonal frequencies (Fig. 4b and Supplementary Fig. 11), with some cases showing only one or two frequency modes (Fig. 4b), indicating a smaller number of clonal genotypes, whereas other tumours exhibit multiple clonal frequency modes, indicating more extensive clonal evolution. Consistent with early 'driver gene' status, mutations in known tumour suppressors such as *p53* tend to occur in the highest clonal frequency group in most tumours. However, in some cases (for example, SA219, SA236; Fig. 4b, Supplementary Fig. 11) *p53* resides in lower-abundance clonal frequency groups (Supplementary Fig. 12 and Fig. 3a), suggesting that it was not the founding event. Although the number of clonal frequency modes tends to increase with the number of mutations, the relationship is not strictly linear (Fig. 4c). To determine whether basal and non-basal cancers differ in their clonality, we compared the distribution of clonal modes (clusters) by case and as an overall distribution, and note that basal TNBCs have more clonal frequency modes than non-basal TNBCs (Fig. 4c). Both of these distributions emphasize a key observation; namely, that at the time of diagnosis TNBCs already display a widely varying clonal evolution that mirrors the variation in mutational evolution.

Finally, we asked where key pathways appear in the distribution of clonal frequency groups. We examined the clonal frequency of genes in each pathway and ascertained if there was a deviation away from the distribution of clonal frequency for all mutations. As expected, pathways involving *p53* and *PIK3CA* showed significantly skewed distributions (Wilcoxon, $q < 0.01$; Fig. 3b and Supplementary Fig. 12) towards higher clonal frequencies, consistent with their roles in early tumorigenesis (Fig. 3a and Supplementary Table 17). Intriguingly, pathways with cytoskeletal genes such as myosins, laminins, collagens and integrins tend to have lower median clonal frequencies, suggesting that somatic mutations in these genes are acquired much later (Fig. 3b). Notably, the median clonal frequency for Reactome pathway 'p53

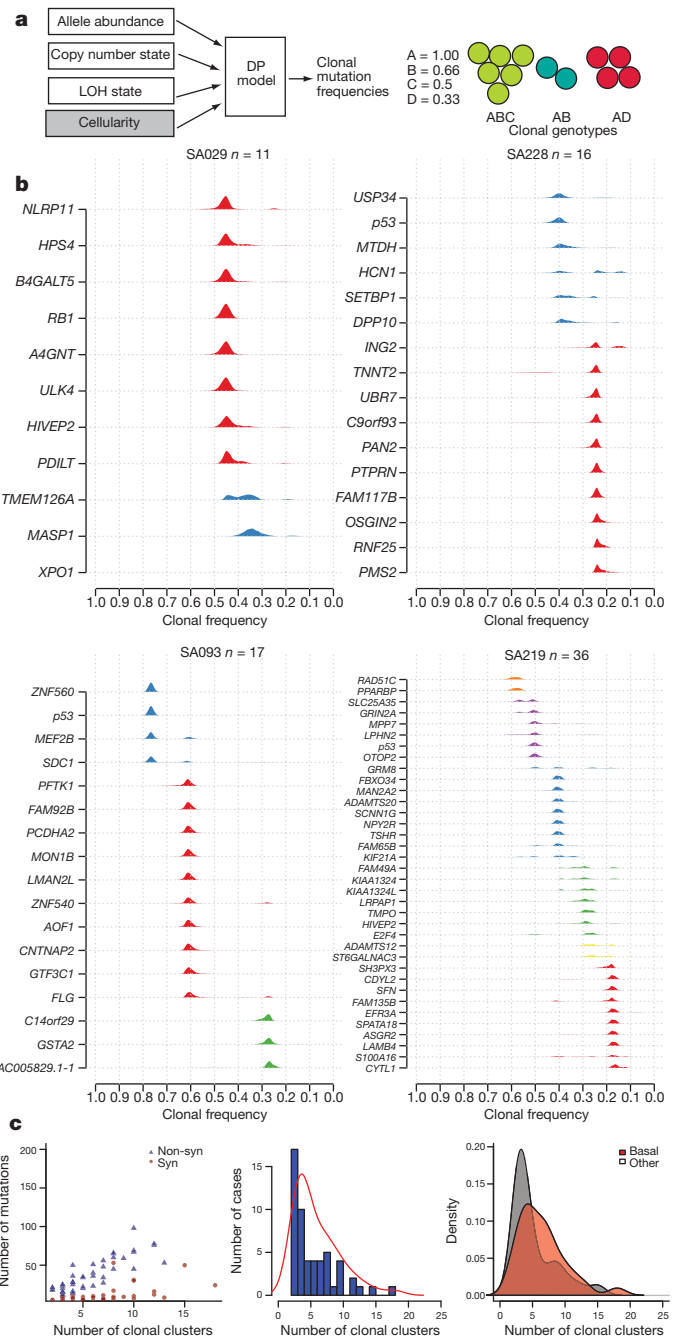


Figure 4 | Clonal evolution in TNBC. **a**, Schematic representation of integration of CNA, LOH, allelic abundance measurements and normal cell contamination for clonal frequency estimation using a Dirichlet process (DP) model (left). Example of a mixture of three clonal genotypes composed of four mutations (A, B, C, D) and their resulting clonal frequencies. **b**, Estimated clonal frequencies for four cases are shown as the distribution of posterior probabilities from the pycclone model (Supplementary Methods). Clonal frequency distributions are coloured by their frequency group membership. **c**, Left, relationship of mutation abundance (synonymous (Syn) and non-synonymous (Non-syn)) and the inferred number of clonal clusters. Middle, distribution and kernel density (red line) of the number of inferred clonal clusters over 54 TNBCs. Right, kernel density distribution of clonal clusters for basal (red) and non-basal (grey) tumours.

pathway feedback loops', including 46 mutations in *ATM*, *ATR*, *NRAS*, *PIK3CA*, *PTEN*, *SIAH1* and *p53*, was 73% (Wilcoxon, $q = 0.0007$), whereas 'integrin cell surface interactions', including 23 mutations in integrin, laminin and collagen genes, had a median clonal frequency of 42% (Wilcoxon, $q = 0.9569$).

Primary TNBCs are still treated as if they were a single disease entity, yet it is clear they do not behave as a single entity in response to current therapies. Here we show for the first time, using next-generation sequencing mutational profiling methods, that treatment-naïve TNBCs display a complete spectrum of mutational and clonal evolution, with some patients' tumours showing only a few somatic coding sequence point mutations with a limited number of molecular pathways implicated, whereas other patients' tumours exhibit considerable additional mutational involvement. Moreover, the clonal heterogeneity of these cancers is also a continuum, with some patients presenting with low-clonality cancers and other cases exhibiting more extensive clonal evolution at diagnosis. In this respect, the basal expression subtype of TNBCs also tends to show higher clonality at diagnosis, although the relationship is not exact.

In clonally evolving tumours, identification of genes by single gene mutation frequency measurements will probably favour early driver genes, because the subsequent involvement of multiple additional pathways during tumour progression is unlikely to be observed as a frequent single gene mutation. The clonality analysis emphasizes this point: known drivers such as *p53*, *PIK3CA* and *PTEN* have among the highest clonal frequencies, whereas mutations in cell shape/motility and ECM-signalling genes appear in the lower clonal frequency groups, distributed over many genes. Although *p53* somatic mutations are clearly early events, the clonal frequencies observed in some TNBC suggest that they are not always the first event, raising a question about what drives early clonal expansion in some of these cancers. Our findings suggest that each TNBC at the time of primary diagnosis may be at a very different phase of molecular progression, with possible implications for approaches to the biology of low clonality versus high clonality primary tumours.

METHODS SUMMARY

The genomes and transcriptomes of 104 TNBCs were profiled with Affymetrix SNP6.0 arrays (all cases), RNA-seq (80 cases; Illumina GAI), and whole exome/genome sequencing (65 cases; tumour and normal DNA). Exomes were obtained using Agilent's Human All Exon SureSelect Target Enrichment System v.1 followed by Illumina GAI sequencing, and whole genomes were sequenced using Life Technologies SOLiD system. Data were analysed using computational approaches to detect somatic SNVs^{4,5}, indels, copy number alterations, gene fusions and gene expression patterns. Predictions were then validated using orthogonal experimental assays, including targeted ultra-deep amplicon sequencing of SNVs to ~20,000× redundancy. We determined single genes under selection using a statistical approach that considers patient-specific background mutation and transition/transversion rates. Mutations predicted to alter transcriptional profiles were determined using an integrated bipartite graph-based method (driverNet) that associates genomic aberrations with outlying expression patterns informed by pre-defined pathway gene sets. Disrupted pathways were determined using the Reactome FI Cytoscape plugin. Clonal analysis was performed (cases with >10 mutations) using a Dirichlet process statistical model that simultaneously estimates clonal frequencies and mutation genotype given deeply sequenced somatic SNVs and copy number estimates. Experimental assays and analytical methodology are detailed in the Supplementary Information.

Received 16 June 2011; accepted 15 February 2012.

Published online 4 April 2012.

- Blows, F. M. *et al.* Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med.* **7**, e1000279 (2010).
- Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
- Sørbye, T. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
- Roth, A. *et al.* JointSNVmix: A probabilistic model for accurate detection of somatic mutations in normal/tumour paired next generation sequencing data. *Bioinformatics* <http://dx.doi.org/10.1093/bioinformatics/bts053> (27 January 2012).
- Ding, J. *et al.* Feature-based classifiers for somatic mutation detection in tumour—normal paired sequencing data. *Bioinformatics* **28**, 167–175 (2012).
- Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069–1075 (2008).
- Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* <http://dx.doi.org/10.1038/nature10983> (this issue).
- Poulogiannis, G. *et al.* *PARK2* deletions occur frequently in sporadic colorectal cancer and accelerate adenoma development in *Apc* mutant mice. *Proc. Natl Acad. Sci. USA* **107**, 15145–15150 (2010).
- Bignell, G. R. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463**, 893–898 (2010).
- Stephens, P. J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462**, 1005–1010 (2009).
- Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298–303 (2011).
- Chicas, A. *et al.* Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence. *Cancer Cell* **17**, 376–387 (2010).
- Herschkowitz, J. I., He, X., Fan, C. & Perou, C. M. The functional loss of the retinoblastoma tumor suppressor is a common event in basal-like and luminal B breast carcinomas. *Breast Cancer Res.* **10**, R75 (2008).
- Langerød, A. *et al.* TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res.* **9**, R30 (2007).
- Børresen-Dale, A.-L. TP53 and breast cancer. *Hum. Mutat.* **21**, 292–300 (2003).
- Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
- Youn, A. & Simon, R. Identifying cancer driver genes in tumor genome sequencing studies. *Bioinformatics* **27**, 175–181 (2011).
- Agrawal, N. *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* **333**, 1154–1157 (2011).
- Stransky, N. *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
- Wu, G., Feng, X. & Stein, L. A human functional protein interaction network and its application to cancer data analysis. *Genome Biol.* **11**, R53 (2010).
- Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
- Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
- Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The support of the BC Cancer Agency Tumour Bank, CBCF Breast Tumour Bank Alberta and the Addenbrookes Tumour bank (supported by NIHR and ECMC) is acknowledged. Technical support is acknowledged from the Centre for Translational Genomics, the Michael Smith Genome Sciences Centre technical group, the BCCA Flow Cytometry Core Facility in the Terry Fox Laboratory and the Cancer Research UK Cambridge Research Institute. Supported by the BC Cancer Foundation, US Department of Defense CDMRP program, Canadian Breast Cancer Foundation (BC Yukon) (to S.A. and S.S.), Michael Smith Foundation for Health Research (to S.S.), US National Institutes of Health (NIH) Roadmap Epigenomics Program, NIH grant 5U01ES017154-02 (to M.H., M.A.M., J.C. and T.T.), Cancer Research UK (to C. Caldas and P.D.P.) and the National Institute of General Medical Sciences (R01GM084875 to W.W.W.), the Canadian Breast Cancer Research Alliance and the Canadian Cancer Society (to S.A. and C.E.). We thank B. Reva, Y. Antipin and C. Sander (Memorial Sloan Kettering Cancer Center) for assistance with MutationAssessor, and G. Wu (Ontario Institute for Cancer Research) for assistance with Reactome.

Author Contributions S.A., S.P.S., C. Caldas and M.A.M. designed and implemented the research plan and wrote the manuscript. S.P.S., A.R., R. Goya, G. Ha, J.D., G. Haffari, A. Bashashati, A. McPherson, K.S., A.C., R. Giuliany, A.H.-M., J.R., D.L., I.B., R.V., S.W.C., M.G., I.M.M., S.J., C. Curtis, O.M.R., P.D.P., V.B. and W.W.W. conducted bioinformatic analyses of the data and/or gave advice on analytic methodology. G.T. conducted histopathological review and immunohistochemistry. A.O., Y.Z., G.T., K.T., L.M.P., J.K., A.B., D.Y., A.T., N.D., T.Z., S.-F.C., K.M. and M.H. conducted sequencing or experimental validation of somatic aberrations. D.Y., A. Moradian, S.-W.G.C. and G.B.M. conducted proteome validation of splicing. P.W., K.G., S.C., S.-F.C., G.T., J.M., C. Caldas, P.D.P. and D.H. collected and interpreted clinical data. S.D., J.F.C., T.T., M.S., P.G. and C.J.E. contributed materials or reagents. K.H., V.T., T.H., M.H. and M.A.M. generated sequence data.

Author Information Aligned exome/genome sequence data, RNA-seq data and Affymetrix SNP6.0 data sets are available at the European Genome-phenome Archive (<http://www.ebi.ac.uk/ega/>) under study accession number EGAS00001000132. Normal reference RNA-seq datasets are available at the NCBI Short Read Archive (<http://www.ncbi.nlm.nih.gov/Traces>) under study accession number SRP000930. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.A. (saparicio@bccrc.ca), C. Caldas (carlos.caldas@cancer.org.uk), S.P.S. (sshah@bccrc.ca) or M.M. (mmarra@bcgsc.ca).

An RNA interference screen uncovers a new molecule in stem cell self-renewal and long-term regeneration

Ting Chen¹, Evan Heller¹, Slobodan Beronja¹, Naoki Oshimori¹, Nicole Stokes¹ & Elaine Fuchs¹

Adult stem cells sustain tissue maintenance and regeneration throughout the lifetime of an animal^{1,2}. These cells often reside in specific signalling niches that orchestrate the stem cell's balancing act between quiescence and cell-cycle re-entry based on the demand for tissue regeneration²⁻⁴. How stem cells maintain their capacity to replenish themselves after tissue regeneration is poorly understood. Here we use RNA-interference-based loss-of-function screening as a powerful approach to uncover transcriptional regulators that govern the self-renewal capacity and regenerative potential of stem cells. Hair follicle stem cells provide an ideal model. These cells have been purified and characterized from their native niche *in vivo* and, in contrast to their rapidly dividing progeny, they can be maintained and passaged long-term *in vitro*⁵⁻⁷. Focusing on the nuclear proteins and/or transcription factors that are enriched in stem cells compared with their progeny^{5,6}, we screened ~2,000 short hairpin RNAs for their effect on long-term, but not short-term, stem cell self-renewal *in vitro*. To address the physiological relevance of our findings, we selected one candidate that was uncovered in the screen: TBX1. This transcription factor is expressed in many tissues but has not been studied in the context of stem cell biology. By conditionally ablating *Tbx1* *in vivo*, we showed that during homeostasis, tissue regeneration occurs normally but is markedly delayed. We then devised an *in vivo* assay for stem cell replenishment and found that when challenged with repetitive rounds of regeneration, the *Tbx1*-deficient stem cell niche becomes progressively depleted. Addressing the mechanism of TBX1 action, we discovered that TBX1 acts as an intrinsic rheostat of BMP signalling: it is a gatekeeper that governs the transition between stem cell quiescence and proliferation in hair follicles. Our results validate the RNA interference screen and underscore its power in unearthing new molecules that govern stem cell self-renewal and tissue-regenerative potential.

Stem cell self-renewal is the process by which stem cells proliferate and generate more stem cells. This process requires control of the cell cycle and maintenance of the undifferentiated state. Embryonic stem cells are refractory to most proliferation checkpoints⁴, and they typically promote self-renewal by suppressing differentiation⁸. By contrast, the few established regulators of self-renewal in adult stem cells function by regulating cell-cycle progression⁹⁻¹². A prerequisite to unlocking the key to regenerative medicine is to dissect the complex mechanisms governing stem cell self-renewal in adult tissues.

With their enormous capacity for tissue regeneration, hair follicles offer an ideal system to explore these mechanisms. Hair follicle stem cells (HF-SCs) become activated early in each hair growth cycle, when a few of these cells exit their niche (called the bulge) to generate a new hair follicle. The differentiation of stem cells into lineage-restricted progeny probably requires micro-environmental stimuli that are not present in the stem cell niche, because this process happens gradually along the follicle outer root sheath¹³⁻¹⁵. The stepwise process culminates at the base of the mature follicle, where committed transit-amplifying matrix cells differentiate into the six lineages that are involved in hair production.

Following their activation, stem cells within the bulge and its vicinity (the upper outer root sheath, which becomes the bulge in the next cycle) briefly self-renew, replenishing the expended stem cells and ensuring long-term tissue regeneration¹³⁻¹⁵. Niche HF-SCs also proliferate following injury and repair wounds¹³⁻¹⁵. Another feature that distinguishes HF-SCs from their committed progeny is their ability to be propagated for at least five passages *in vitro*, reflecting their capacity for long-term proliferative potential⁶.

In the current study, we surmised that there might be two sources for finding intrinsic factors responsible for maintaining 'stemness' inside and outside the stem cell niche: self-renewal factors that have been identified in other stem cell studies; and nuclear proteins that we found to be enriched twofold in stem cells relative to their transit-amplifying progeny (Supplementary Fig. 1a, b). Focusing on about 400 such candidates, we devised an *in vitro* RNA interference (RNAi) screen for long-term versus short-term self-renewal (Fig. 1a). By choosing genes whose expression was enriched in stem cells relative to committed proliferative progeny, this pool of candidates should not contain housekeeping genes and general proliferation-associated genes. However, if short hairpin RNAs (shRNAs) target a gene that is essential for long-term but not short-term self-renewal, then cells expressing this gene should persist during early passages but then decrease in number or disappear with sequential passaging. Operating on this premise, we transduced purified primary HF-SCs in triplicate with a lentiviral pool encoding control (scramble) shRNAs and a pool of 2,035 candidate shRNAs (about five per gene) such that, on average, each stem cell expressed a single shRNA (Supplementary Fig. 1c). The transduced stem cells were cultured and, at 24 h and following each passage, shRNAs were amplified from the surviving cells and subjected to high-throughput sequencing.

Data are shown for passage 1 (P-1) and P-5 (Fig. 1b, c, Supplementary Figs 2 and 3a, and Supplementary Tables 1 and 2). More than 96% of the initial shRNAs were detected at 24 h after transduction, and these shRNAs were used as a reference for changes in shRNA representation. Consistent with our strategic exclusion of housekeeping genes and general proliferation-associated genes, most cells that harboured shRNAs survived the first passage. By contrast, after five passages, many shRNAs were depleted or enriched, suggesting that the transduced cells had different long-term proliferative potentials. Using unsupervised hierarchical clustering, triplicates of individually transduced and passaged cells behaved strikingly similarly, suggesting that these changes reflected bona fide alterations in stem cell character.

Parallel screens with fibroblasts weeded out shRNAs corresponding to cell-survival genes such as *Bcl2*, which were selected against after five passages in both HF-SCs and fibroblasts (Fig. 1c, Supplementary Fig. 3b and Supplementary Table 3). Our refined short list of self-renewal candidates contained those whose cognates all showed similar trends and for which two or more shRNAs per gene displayed specific changes in P-5 stem cell cultures but not in P-1 stem cell cultures or in P-5 fibroblasts (Supplementary Fig. 2 and Supplementary Table 1). Category I shRNAs (Fig. 1d) were maintained in P-1 stem cell cultures but were underrepresented by more than 90% at P-5, meeting the criteria for an shRNA that suppresses a long-term self-renewal gene.

¹Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10065, USA.

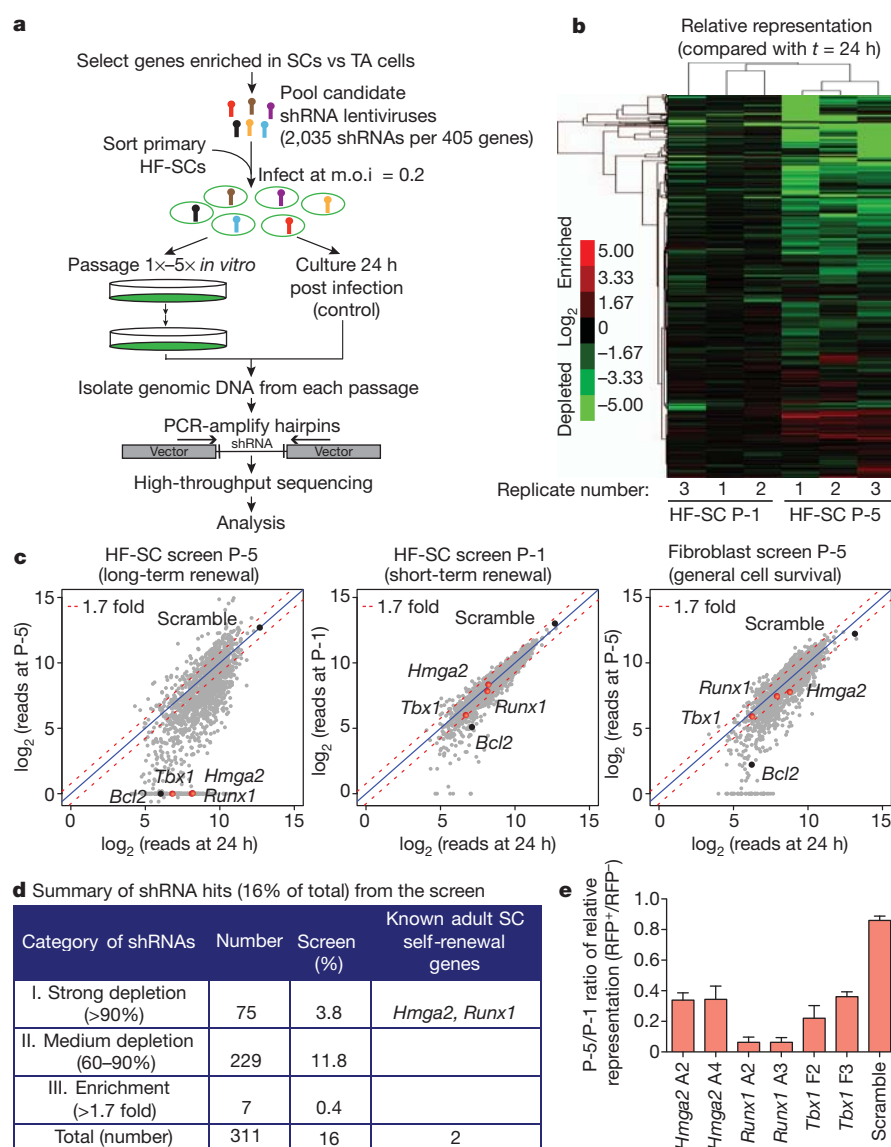


Figure 1 | *In vitro* RNAi screen for genes involved in stem cell long-term self-renewal. **a**, *In vitro* RNAi screening strategy. **b**, Unsupervised hierarchical clustering of screening replicates. **c**, Scatter plots of normalized reads per shRNA between 24 h post infection and after one passage (P-1) or five passages (P-5) for HF-SCs (left and centre) and after five passages for fibroblasts (right). The data shown are from one replicate of each screening, highlighting the genes whose corresponding shRNAs were specifically depleted in the long-term passaging of stem cells (red dots) and control genes (black dots). The blue line is the diagonal line with a ratio of 1.0. The red dashed line shows the cut off for 1.7-fold change. **d**, Screening statistics. **e**, Progressive selection against hairpins that target putative long-term self-renewal genes. Data are presented as mean \pm s.d.; $n = 3$. m.o.i., multiplicity of infection; RFP, red fluorescent protein; SC, stem cell; TA, transit-amplifying; vs, versus.

Representing only 3.8% of the initial pool, category I included shRNAs targeting *Hmga2*, which is required for neural stem cell self-renewal¹⁰, and *Runx1*, which promotes HF-SC proliferation¹⁶.

Real-time PCR (rtPCR) of transcripts targeted by six of the most effective category I shRNAs confirmed that each shRNA blocked the expression of its intended target (Supplementary Fig. 4). Moreover, stem cells that were individually transduced with *Hmga2*, *Runx1* or *Tbx1* shRNA were progressively selected against over time (Fig. 1e). The transcription factor TBX1 was particularly intriguing because it has been implicated in tissue formation in other organs^{17,18}. We selected it as our model for *in vivo* testing of the functional relevance of our RNAi screen.

rtPCR and epigenetic chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) analyses¹⁹ of purified hair follicle populations revealed that *Tbx1* was transcribed at higher levels in stem cells than in any of their progeny (Fig. 2a, b and Supplementary Fig. 5a). *In vivo*, the developmental expression of TBX1 protein most closely resembled that of two essential HF-SC transcription factors, SOX9 and LHX2 (Fig. 2c, d). The adult pattern of expression resembled that of CD34, which is a cell surface marker of HF-SCs (Fig. 2e, f). Nuclear TBX1 was not detected in self-renewing transit-amplifying cells or in terminally differentiating cells (Fig. 2g). Notably, in contrast to some other HF-SC transcription factors, TBX1 was also maintained in stem cells in long-term cultures.

To evaluate TBX1 function *in vivo*, we conditionally targeted *Tbx1* (*Tbx1*-cKO) in the skin epithelium of embryonic day 15.5 mice^{20,21}. *Tbx1*-cKO mice were viable, and hair follicle morphogenesis appeared to be normal (Supplementary Fig. 5b, c). We tested for possible defects in stemness by analysing tissue regeneration during the normal hair cycle (Supplementary Fig. 6). For this purpose, same-sex littermates were shaved at the normal onset of the first hair cycle (postnatal day 21, P21) and the second hair cycle (P60). In both cases, *Tbx1*-cKO hair follicles remained quiescent longer than normal, but they eventually cycled. Maturation/differentiation was unaffected, as shown by the development of normal hair types and lengths.

Self-renewal occurs briefly after anagen onset, replenishing the stem cells that are used during initiation¹⁵. To challenge stem cells to sustain long-term tissue regenerative potential, we repeatedly depilated the hair coat, a process that removes the old hair along with tightly adhering niche signalling cells that maintain stem cell quiescence^{14,15} (Fig. 3a). After depilation, more than 80% of the stem cells remained viable in their niche, where they became activated to enter a new hair cycle (Supplementary Fig. 7).

Wild-type (WT) HF-SCs survived each round of depilation-induced hair regeneration, indicating the robust ability to sustain self-renewal and long-term tissue regeneration. By contrast, after five rounds, the *Tbx1*-cKO stem cell numbers had declined by more than 70% ($P < 0.001$) (Fig. 3b, c). Their steady depletion was accompanied

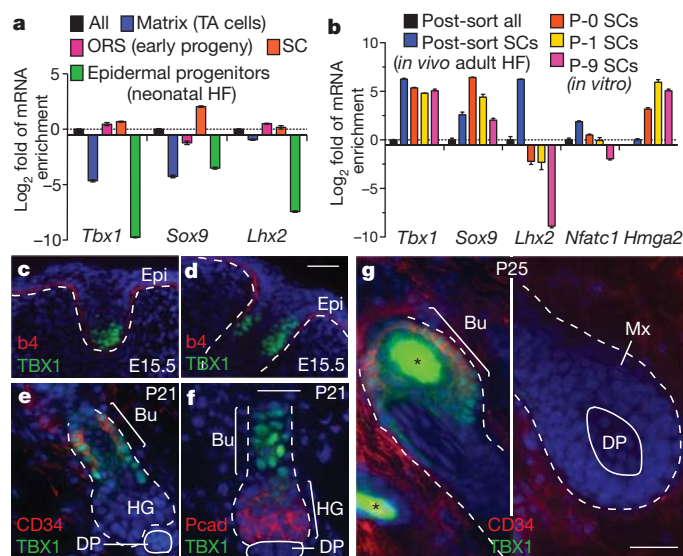


Figure 2 | The transcription factor TBX1 is highly enriched in stem cells *in vivo* and *in vitro*. **a, b,** *Tbx1* expression in HF-SCs, shown compared with established HF-SC transcription factor genes, as determined by rtPCR. The mRNAs were isolated from fluorescence-activated cell sorting (FACS)-purified stem cells or other populations, as indicated. Data are presented as mean \pm s.d.; $n = 3$. **c–g,** Nuclear TBX1 in the skin is restricted to developing and postnatal HF-SCs. The back skins of mice were processed for immunofluorescence at embryonic day 15.5 (E15.5), postnatal day 21 (P21) and P25 (P25). The white dashed line indicates the dermal–epithelial boundary, and the asterisk denotes hair shaft autofluorescence. Nuclei are shown in blue. Scale bars, 30 μ m. **b4,** β_4 -integrin; **Bu,** bulge stem cells; **DP,** dermal papilla; **Epi,** epidermis; **HF,** hair follicle; **HG,** hair germ (early stem cell progeny); **Mx,** matrix (committed TA progeny); **ORS,** outer root sheath; **Pcad,** P-cadherin.

by a thinning of the hair coat and a reduction in hair follicle density (Fig. 3d, e and Supplementary Fig. 8). Histological analysis revealed that many *Tbx1*-cKO hair follicles were dormant and had lost their

stem cell niche, retaining only sebaceous glands and dermal papillae. However, the few hair follicles that cycled appeared morphologically normal, reflecting the presence of active stem cell niches (Fig. 3e).

Similar results were obtained when hair cycles were monitored during natural ageing. Although the intervals between the hair cycles were longer than those in the depilation procedure, the *Tbx1*-cKO stem cell niche residents had declined by about 30% after 1 year (Fig. 3f). Thus, *Tbx1*-null stem cells seem to be specifically impaired in their long-term ability to replenish their niche during normal and depilation-induced tissue regeneration.

We used 5-bromodeoxyuridine (BrdU) incorporation to define the brief window of bulge stem cell proliferation that occurs following depilation. WT stem cell proliferation peaked at day 3 after depilation, and the cells returned to quiescence by day 7. *Tbx1*-cKO stem cells also proliferated but to a lesser extent during this time period (Fig. 4a). Within 2 to 3 days of depilation, only about 25% of *Tbx1*-null stem cells were BrdU-positive, whereas about 70% of WT stem cells were BrdU-positive (Fig. 4a). This proliferative decrease was verified by DNA-content-based cell-cycle analysis of purified stem cells, which showed that the decrease was accompanied by fewer stem cells being present in the niche (Fig. 4b). As discussed earlier, most hair follicles eventually produced hairs of WT length, reflecting an otherwise normal lineage program.

We observed a similar trend when we monitored the normal hair cycle. In this case, fewer stem cells were expended than during depilation; thus, the demand for self-renewal was lower, as reflected by the natural bulge niche having a lower proliferative activity than the depilation-induced WT bulge niche. Consistent with a role for TBX1 in HF-SC self-renewal, the overall proliferative activity within the *Tbx1*-null niche was less than in the WT niche, and that in the natural niche was less than in the depilation-induced niche (Supplementary Fig. 9).

To understand how these differences arise, we transcriptionally profiled messenger RNAs that were isolated from purified HF-SCs 2 days after depilation (Fig. 4c, Supplementary Fig. 10 and Supplementary Table 4). Bioinformatic analysis using the Database for Annotation,

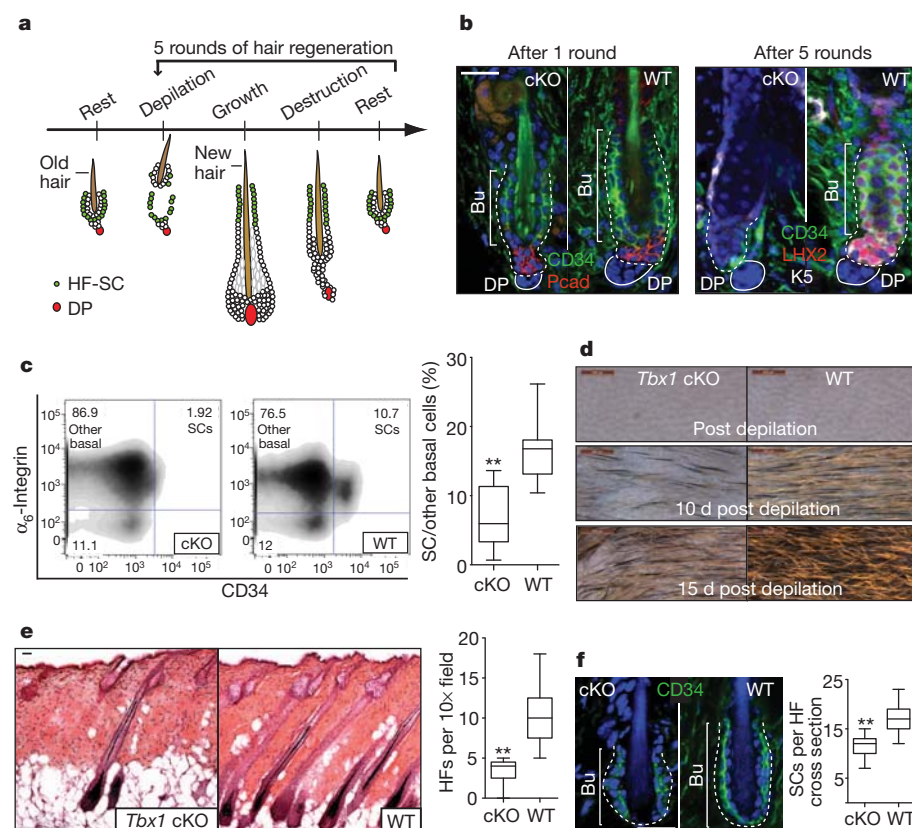


Figure 3 | *Tbx1*-null stem cells fail in an *in vivo* assay for stem cell self-renewal and long-term tissue regeneration. **a,** Schematic of the *in vivo* long-term hair regeneration assay. **b,** Decline in stem cell number with sequential rounds of hair regeneration in *Tbx1*-cKO hair follicles. **c,** FACS quantifications of stem cells after five rounds of hair regeneration; $n = 11$ for WT, $n = 12$ for cKO. **d,** Close up of skin surface in the fifth cycle of hair regeneration. **e,** Representative haematoxylin and eosin images and quantifications of follicle densities in the fifth cycle of hair regeneration; $n = 4$. **f,** Decline in stem cell number during normal ageing (1 year old) in *Tbx1*-cKO skin; $n = 5$. **c, e, f,** Box-and-whisker plots: mid-line, median; box, 25th to 75th percentiles; and whiskers, minimum and maximum. **, $P < 0.001$. **b, d, e, f,** The white dashed line indicates the dermal–epithelial boundary, and nuclei are shown in blue. Scale bars, 500 μ m (**d**) and 30 μ m (**b, e, f**).

Visualization and Integrated Discovery (DAVID) functional gene annotation tool uncovered mostly cell-cycle regulators in the 123 genes that were downregulated by 1.8 fold or more in *Tbx1*-null stem cells compared with WT stem cells. The 188 genes that were upregulated by 1.8 fold or more were enriched for genes implicated in bone morphogenetic protein (BMP) signalling (Fig. 4c, red), including *Bmp2*, which encodes a secreted ligand for BMP receptors, and *Id* genes, which are major targets of the transcriptional effectors of BMPs, namely SMAD4 in complex with phosphorylated SMAD1 (pSMAD1), pSMAD5 or pSMAD8 (denoted pSMAD1/5/8). Using rtPCR, *Tbx1*-cKO HF-SCs upregulated *Id1* and *Id2* (refs 22, 23) (Fig. 4d).

In cardiomyocytes, TBX1 seems to suppress BMP signalling by competitively interfering with SMAD4 for pSMAD1/5/8 binding²⁴. Consistent with this idea, overexpression of a TBX1–green fluorescent protein (GFP) fusion protein in WT keratinocytes significantly suppressed BMP-induced *Id1* transcription *in vitro* (Fig. 4e). Similar effects, albeit slower and weaker, were observed for *Id2*.

Transgenic overactivation of the BMP circuitry results in hair coat thinning with age²⁵. If TBX1-deficient HF-SCs have a heightened sensitivity to BMP signalling, then BMP inhibitors might ameliorate the proliferative defect. We tested this hypothesis by plucking hair follicles from mice and then injecting them intradermally with beads soaked in the BMP antagonist noggin^{26,27}. Within 3 days, *Tbx1*-cKO HF-SC proliferation had been restored to near WT levels (Fig. 4f,

$P < 0.001$, and Supplementary Fig. 11). As reflected by the bulge size, the expanded HF-SC pool was sustained throughout the hair cycle. However, additional treatments with noggin were necessary to maintain the stem cell pool through multiple rounds of depilation-induced hair regeneration. Interestingly, the self-renewal of TBX1-deficient stem cells was also elevated *in vitro* when BMP signalling was impaired by ablation of the BMP receptor BMPRI1A (Fig. 4g). Together, these findings are consistent with a BMP-induced proliferative defect in *Tbx1*-null stem cells.

Given these inverse links between TBX1 and BMP signalling, it seemed paradoxical that *Smad1* shRNAs surfaced in our self-renewal screen (Supplementary Table 1). Further analyses revealed that even though these shRNAs depleted *Smad1* transcripts and SMAD1 protein, the SMAD1/5/8 target genes *Id1*, *Id2* and *Id3* were still expressed. Moreover, the transduced cells still responded to BMP signalling, as judged by reporter assays (Supplementary Fig. 12).

HF-SCs reside in a WNT-restricted, BMP-high micro-environment^{28,29} in which they must self-renew to replenish the stem cell pool. Therefore, HF-SCs must have an intrinsic mechanism to lower the BMP signalling threshold, and this mechanism fails to occur in the absence of TBX1. Entering the hair cycle also necessitates decreased BMP signalling; however, in this case, the proliferation is fuelled by early progeny (hair germ) with naturally low TBX1 levels that are present at the bulge base^{6,30}. Because, paradoxically, hair cycle

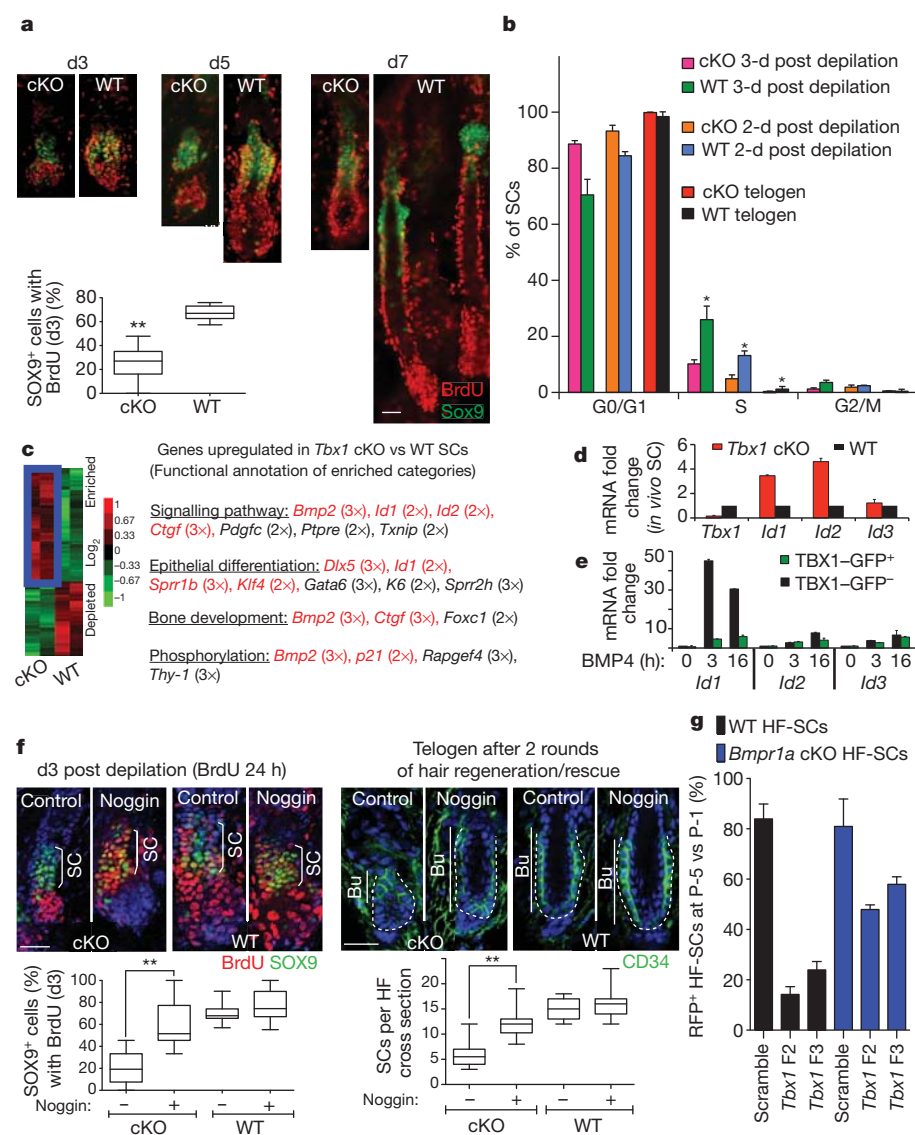


Figure 4 | TBX1 controls stem cell proliferation in part by fine-tuning the response to BMP signalling.

a, Stem cell proliferative status post depilation, as assessed by BrdU incorporation (with BrdU added 24 h before analysis); $n = 3-5$. **b**, Cell-cycle analysis of FACS-isolated HF-SCs; $n = 3-5$; *, with respect to cKO. **c**, Gene expression analysis of HF-SCs. Genes upregulated in HF-SCs from *Tbx1*-cKO mice compared with WT HF-SCs are delineated in blue. Functional annotations for the enriched categories are shown. **d**, rtPCR of BMP targets in HF-SCs. **e**, rtPCR of *Id1*, *Id2* and *Id3* post BMP4 treatment with or without TBX1-GFP. **f**, Transient inhibition of BMP signalling by noggin *in vivo* rescues stem cell numbers in *Tbx1*-cKO mice; $n = 3$ and ~ 50 hair follicles each. The white dashed line indicates the dermal-epithelial boundary, and nuclei are shown in blue. **g**, Long-term passage of *Tbx1*-shRNA-transduced HF-SCs; $n = 3$. **b**, **d**, **g**, Data are presented as mean \pm s.d. *, $P < 0.01$. **a**, **f**, Box-and-whisker plots: mid-line, median; box, 25th to 75th percentiles; and whiskers, minimum and maximum. **, $P < 0.001$. Scale bars, 30 μ m.

initiation was delayed in *Tbx1*-null mice, we surmise that the hair germ may be negatively influenced by BMP2 or other local effectors that are secreted by *Tbx1*-null HF-SCs.

In summary, we have discovered that TBX1 functions in the replenishment of HF-SCs during tissue regeneration. Our RNAi screens excluded roles for TBX1 in cell cycling, housekeeping and survival. The finding that once initiated, *Tbx1*-null hair cycles have a normal progression also ruled out roles for TBX1 in cell-fate determination or lineage progression. Instead, the TBX1 defect seems to be rooted in diminished stem cell self-renewal, coupled with enhanced sensitization to intrinsic BMP signalling. Together, these result in progressive HF-SC depletion and thinning of the hair coat. Although the effects of TBX1 are likely to be more complex than we have shown, its ability to intrinsically control these features poises it in the middle of a balancing act with the micro-environment to control stem cell behaviour in tissue homeostasis.

METHODS SUMMARY

RNAi screen. Candidate gene selection was based on previously published microarray analyses^{5,6}. HF-SCs isolated by fluorescence-activated cell sorting (FACS) were infected with an shRNA library (carried by TRC lentiviruses) targeting a set of ~400 candidate genes (2,035 shRNAs, with approximately 5 shRNAs per gene) to a final infection rate of approximately 20% (that is, about one virus per five HF-SCs). Twenty-four hours after infection, half of the infected HF-SCs were collected, and the other half were plated onto mitomycin-C-treated fibroblast feeder layers. Each week, nearly confluent cultures were trypsinized and replated (one passage). At each passage, a fraction of the cells were processed for genomic DNA isolation. Primers, including adaptors for Solexa sequencing, were used to amplify shRNA-encoding sequences from genomic DNA. Following the PCR amplification of shRNAs, sequencing was performed on an Illumina/Solexa Genome Analyzer II according to the manufacturer's protocols. Analyses and plots of DNA sequencing data were performed in the R statistical environment. All shRNA identities, as well as primary screening data, are listed in Supplementary Tables 1–3.

Animal studies. *Tbx1*^{fl/fl} mice were obtained from A. Baldini. To create conditional knockout mice, we mated hemizygous *K14-Cre* (CD1) mice with homozygous *Tbx1*^{fl/fl} (C57BL/6) mice; F₁ *K14-Cre/Tbx1*^{fl/+} (CD1/C57BL/6) progeny were subsequently bred with homozygous *Tbx1*^{fl/fl} mice to generate *K14-Cre;Tbx1*^{fl/fl} mice at a 25% Mendelian frequency. Depilation of mid-dorsal hair follicles was achieved on anaesthetized mice to provide a proliferative stimulus and to synchronize a population of anagen hair follicles. The mid-dorsum was coated with molten wax, which was peeled off after hardening. BrdU (Sigma-Aldrich) pulse experiments involved intraperitoneal injections (50 µg BrdU per g body weight) twice a day. After BrdU pulses for the indicated times, animals were killed. The skins were embedded in OCT compound, frozen in dry ice, sectioned using a cryostat (Richard-Allan Scientific) and stained for immunofluorescence using antibodies specific for BrdU.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 September 2011; accepted 9 February 2012.

Published online 4 April 2012.

- Li, L. & Clevers, H. Coexistence of quiescent and active adult stem cells in mammals. *Science* **327**, 542–545 (2010).
- Fuchs, E. The tortoise and the hair: slow-cycling cells in the stem cell race. *Cell* **137**, 811–819 (2009).
- Zon, L. I. Intrinsic and extrinsic control of haematopoietic stem-cell self-renewal. *Nature* **453**, 306–313 (2008).
- He, S., Nakada, D. & Morrison, S. J. Mechanisms of stem cell self-renewal. *Annu. Rev. Cell Dev. Biol.* **25**, 377–406 (2009).
- Blanpain, C., Lowry, W. E., Geoghegan, A., Polak, L. & Fuchs, E. Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* **118**, 635–648 (2004).
- Greco, V. *et al.* A two-step mechanism for stem cell activation during hair regeneration. *Cell Stem Cell* **4**, 155–169 (2009).
- Watt, F. M. & Jensen, K. B. Epidermal stem cell diversity and quiescence. *EMBO Mol. Med.* **1**, 260–267 (2009).
- Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. BMP induction of *Id* proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281–292 (2003).
- Molofsky, A. V. *et al.* *Bmi-1* dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962–967 (2003).

- Nishino, J., Kim, I., Chada, K. & Morrison, S. J. Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16^{INK4a} and p19^{Arf} expression. *Cell* **135**, 227–239 (2008).
- Topley, G. I., Okuyama, R., Gonzales, J. G., Conti, C. & Dotto, G. P. p21^{WAF1/Cip1} functions as a suppressor of malignant skin tumor formation and a determinant of keratinocyte stem-cell potential. *Proc. Natl Acad. Sci. USA* **96**, 9089–9094 (1999).
- Kippin, T. E., Martens, D. J. & van der Kooy, D. p21 loss compromises the relative quiescence of forebrain stem cell proliferation leading to exhaustion of their proliferation capacity. *Genes Dev.* **19**, 756–767 (2005).
- Tumbar, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004).
- Ito, M., Kizawa, K., Hamada, K. & Cotsarelis, G. Hair follicle stem cells in the lower bulge form the secondary germ, a biochemically distinct but functionally equivalent progenitor cell population, at the termination of catagen. *Differentiation* **72**, 548–557 (2004).
- Hsu, Y. C., Pasolli, H. A. & Fuchs, E. Dynamics between stem cells, niche, and progeny in the hair follicle. *Cell* **144**, 92–105 (2011).
- Osorio, K. M. *et al.* Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation. *Development* **135**, 1059–1068 (2008).
- Aggarwal, V. S. *et al.* Mesodermal *Tbx1* is required for patterning the proximal mandible in mice. *Dev. Biol.* **344**, 669–681 (2010).
- Chen, L., Fulcoli, F. G., Tang, S. & Baldini, A. *Tbx1* regulates proliferation and differentiation of multipotent heart progenitors. *Circ. Res.* **105**, 842–851 (2009).
- Lien, W. H. *et al.* Genome-wide maps of histone modifications unwind *in vivo* chromatin states of the hair follicle lineage. *Cell Stem Cell* **9**, 219–232 (2011).
- Xu, H. *et al.* *Tbx1* has a dual role in the morphogenesis of the cardiac outflow tract. *Development* **131**, 3217–3227 (2004).
- Vasioukhin, V., Degenstein, L., Wise, B. & Fuchs, E. The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc. Natl Acad. Sci. USA* **96**, 8551–8556 (1999).
- Korchynskyi, O. & ten Dijke, P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the *Id1* promoter. *J. Biol. Chem.* **277**, 4883–4891 (2002).
- Hollnagel, A., Oehlmann, V., Heymer, J., Ruther, U. & Nordheim, A. *Id* genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *J. Biol. Chem.* **274**, 19838–19845 (1999).
- Fulcoli, F. G., Huynh, T., Scambler, P. J. & Baldini, A. *Tbx1* regulates the BMP–Smad1 pathway in a transcription independent manner. *PLoS ONE* **4**, e6049 (2009).
- Blessing, M., Nanney, L. B., King, L. E., Jones, C. M. & Hogan, B. L. Transgenic mice as a model to study the role of TGF-β-related molecules in hair follicles. *Genes Dev.* **7**, 204–215 (1993).
- Botchkarev, V. A. *et al.* Noggin is a mesenchymally derived stimulator of hair-follicle induction. *Nature Cell Biol.* **1**, 158–164 (1999).
- Kulesa, H., Turk, G. & Hogan, B. L. Inhibition of Bmp signaling affects growth and differentiation in the anagen hair follicle. *EMBO J.* **19**, 6664–6674 (2000).
- Kobiak, K., Stokes, N., de la Cruz, J., Polak, L. & Fuchs, E. Loss of a quiescent niche but not follicle stem cells in the absence of bone morphogenetic protein signaling. *Proc. Natl Acad. Sci. USA* **104**, 10063–10068 (2007).
- Andl, T. *et al.* Epithelial *Bmpr1a* regulates differentiation and proliferation in postnatal hair follicles and is essential for tooth development. *Development* **131**, 2257–2268 (2004).
- Oshimori, N. & Fuchs, E. Paracrine TGF-β signaling counterbalances BMP-mediated repression in hair follicle stem cell activation. *Cell Stem Cell* **10**, 63–75 (2012).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to A. Baldini for the *Tbx1*^{fl/fl} mice, the Comparative Biology Centre (AAALAC-accredited) for expert handling and care of the mice, The Rockefeller University FCRC for FACS sorting (supported by NYSTEM funds through NYSDOH contract C023046), S. Dewell and The Rockefeller University Genomic Resource Centre for high-throughput sequencing, The Memorial Sloan Kettering Genomics Core Facility for RNA and microarray processing, Fuchs' lab members Y. Hsu, B. Keyes, X. Wu, D. Devenport and L. Zhang for comments and suggestions, W. H. Lien for providing epigenetic ChIP-seq data for the *Tbx1* gene *in vivo* and P. Janki for assisting in the production of the pooled virus for screening. This work was supported by grants from the National Institutes of Health (R01-AR050452; E.F.), the Empire State Stem Cell (NYSTEM N09G074; E.F.), a New York Stem Cell Foundation-Druckenmiller Fellowship (T.C.) and a NYSTEM Scholar Award (C026722; T.C.). E.F. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions T.C. and E.F. designed the study. T.C. carried out the experiments and analysed the data. E.H. analysed the sequencing data. S.B. contributed to optimizing and testing methods used in the screen. N.O. carried out the *Smad1* shRNA-related experiments. N.S. participated in the experiments involving mouse handling. T.C. and E.F. co-wrote the paper. E.F. supervised the research.

Author Information Microarray data have been deposited in the Gene Expression Omnibus database under accession number GSE35575. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.F. (fuchs@rockefeller.edu).

METHODS

RNAi screen. Candidate gene selection was carried out based on previously published microarray analyses^{5,6}. Fluorescence-activated cell sorting (FACS)-isolated HF-SCs were infected with an shRNA library (carried by TRC lentiviruses) targeting a set of ~400 candidate genes (2,035 shRNAs, with approximately 5 shRNAs per gene for most, if not all, candidates) to a final infection rate of approximately 20% (that is, about one virus per five HF-SCs). Twenty-four hours after infection, half of the infected HF-SCs were collected, and the other half were plated onto mitomycin-C-treated fibroblast feeder layers. Each week, nearly confluent cultures were trypsinized and replated (one passage). At each passage, a fraction of the cells were processed for genomic DNA isolation. Primers, including adaptors for Solexa sequencing, were used to amplify shRNAs from genomic DNA. Following the PCR amplification of hairpins, sequencing was performed on an Illumina/Solexa Genome Analyzer II according to the manufacturer's protocols. Analyses and plots of DNA sequencing data were performed in the R statistical environment³¹. The fold change in shRNA representation after sequential passages *in vitro* was determined by comparing the shRNA representation in each sample to that in the control cell population collected 24 h after infection. The identities of all 'hit' shRNAs, as well as primary fold changes, are listed in Supplementary Tables 1–3.

Fluorescence based competition assay. shRNA-transduced (RFP⁺) HF-SCs (GFP⁺) were mixed with non-transduced HF-SCs and plated onto feeders (GFP⁻). At each passage, the cells were trypsinized and replated, and a fraction of the cells were used to measure the proportion of RFP⁺GFP⁺/RFP⁻GFP⁺ cells using flow cytometry on an LSRII FACS Analyzer (BD Biosciences). Analyses were carried out for five consecutive passages.

RNAi construct and sequences. The RNAi lentiviral constructs were from the RNAi Consortium (TRC) mouse lentiviral library. Scramble shRNA (Ctrl shRNA) that does not target any mouse mRNA was used as a control. Lentiviral vectors were packaged, concentrated and used for infection as previously described³². Sequences of individual shRNAs used in experiments are listed here, together with their symbol in oligoSeq. *Tbx1* F2, CCGGGCTACCGGTATGCTTCCATAC TCGAGTATGGAAGCATACCGGTAGCTTTTGT; *Tbx1* F3, CCGGCTGACC AATAACCTGTGGATCTCGAGATCCAGCAGGTATTGGTCAGTTTTTTGT; *Hmga2* A2, CCGGGCCACAACAAGTCGTTTCCAGAACTCGAGTTCTGAACG ACTTGTGTGGCTTTTTTGT; *Hmga2* A4, CCGGAGACCTAGGAAATGGCCA CAACCTCGAGTTGTGGCCATTTCTAGGTCTTTTTTGT; *Runx1* A2 CCGGG CCCTCCTACCATCTATACTACTCGAGTAGTATAGATGGTAGGAGGGCT TTTTTGT; *Runx1* A3, CCGGGTCTTTTACAAATCCGCCACAACCTCGAGTTGTG CGGATTTGTAAAGACTTTTTTGT; *Smad1* B3, CCGGGCCGAGTAACCTGC GTCCACATCTCGAGATGGTGACGAGTTACTCGGCTTTTTTGT; *Smad1* F1, CCGGCCCATTTGGTTCCAAGCAGAACTCGAGTTCTGCTTGAACCAAAA TGGGTTTTTGT; *Smad1* F2, CCGGACCGTGTATGAACCTACCAAACTCGAGT TTGGTAGTTTCATACACGGTTTTTGT; *Smad1* F4, CCGGTGGTGCTCTATT GTGTACTATCTCGAGATAGTACACAATAGAGCACCATTTTTTGT; and *Smad1* F5, CCGGTCTATTTTCATCCGTGTCTTACTCGAGTAAGACACGGATGAA ATAGGATTTTTTGT.

Mice, depilation and BrdU labelling. *Tbx1*^{fl/fl} mice were obtained from A. Baldini. To create conditional knockout mice, we mated hemizygous *K14-Cre* (CD1) mice with homozygous *Tbx1*^{fl/fl} (C57BL/6) mice; F₁ *K14-Cre/Tbx1*^{fl/+} (CD1/C57BL/6) progeny were subsequently bred with homozygous *Tbx1*^{fl/fl} mice to generate *K14-Cre;Tbx1*^{fl/fl} mice at a 25% Mendelian frequency. Depilation of mid-dorsal hair follicles was achieved on anaesthetized mice to provide a proliferative stimulus and to synchronize a population of anagen hair follicles. The mid-dorsum was coated with molten wax, which was peeled off after hardening. BrdU (Sigma-Aldrich) pulse experiments involved intraperitoneal injections (50 µg BrdU per g body weight) twice a day. After BrdU pulses for the indicated times, animals were killed. The skins were embedded in OCT compound, frozen in dry ice, sectioned using a cryostat (Richard-Allan Scientific) and stained for immunofluorescence using antibodies specific for BrdU. All animals were maintained in an AAALAC-approved animal facility, and procedures were performed with IACUC-approved protocols.

Semi-quantitative RT-PCR. RNAs were isolated from cells using an RNeasy kit (QIAGEN), and DNase treatment was performed to remove genomic DNA. Equal RNA amounts were added to reverse-transcriptase reaction mix (Invitrogen) with oligo(dT)₁₂ as a primer. Semi-quantitative PCR was conducted with a LightCycler system (Roche Diagnostics). Reactions were performed using the indicated primers and template mixed with the LightCycler DNA Master SYBR Green kit and were run for 45 cycles. The specificity of the reactions was determined by subsequent melting curve analysis. LightCycler analysis software was used to remove background fluorescence (noise band). The number of cycles needed to reach the crossing point for each sample was used to calculate the amount of each product using the 2^{-(ΔΔCt)} method. The levels of PCR product were expressed as a function of *Ppib*. The primers were designed to produce a product spanning

exon–intron boundaries. The sequences of the primers were as follows: *Ppib* sense, GTGAGCGCTTCCCAGATGAGA; *Ppib* antisense, TGCCGGAGTCGACAAT GATG; *Sox9* sense, CGGCGGAGGAAGTCGGTGAAGAAC; *Sox9* antisense, GTGGGTGCGGTGCTGCTGATG; *Tbx1* sense, GCTGTGGGACGAGTTCA ATC; *Tbx1* antisense, ACGTGGGGAACATTCGCTCT; *Lhx2* sense, CCAGCTTC GGACAATGAAGT; *Lhx2* antisense, TTTCCTGCCGTAAGAGTTG; *Nfatc1* sense, AACGCCCTGACCACCGATAGCACT; *Nfatc1* antisense, CCCGGCT GCCTTCCGTCTCAT; *Runx1* sense, CTCCGTGCTACCCACTCACT; *Runx1* antisense, ATGACGGTGACCAGAGTGC; *Hmga2* sense, AAGGCAGCAAAA ACAAGAGC; *Hmga2* antisense, CCGTTTTTCTCCAATGGTCT; *Smad1* sense, AACACCAGGCGACATATTGG; *Smad1* antisense, CACTGAGGCATTCCG CATA; *Smad5* sense, GCAGTAACATGATTCCCTCAGACC; *Smad5* antisense, GCGACAGGCTGAACATCTCT; *Smad8* sense, CGGATGAGCTTTGTGAA GG; *Smad8* antisense, GGGTGCTCGTGACATCCT; *Id1* sense, GAGTCTGAAG TCGGGACCAC; *Id1* antisense, TTTCCTCTTGCTCTCTGAA; *Id2* sense, AATGGCCTTTTTGACACGAG; *Id2* antisense, AAAGCAAGCAATCAACA TTCAA; *Id3* sense, GAGGAGCTTTTGCCACGAC; and *Id3* antisense, TGAAGAGGGCTGGGTTAAGA.

Histology and immunofluorescence. Tissues were embedded in OCT compound, and frozen sections were fixed in 4% paraformaldehyde and subjected to immunofluorescence microscopy or haematoxylin and eosin staining as previously described¹. The antibodies (and their dilutions) used were as follows: anti-LHX2 (rabbit, 1:2,500, Fuchs lab), anti-SOX9 (rabbit, 1:1,000, Fuchs lab), anti-P-cadherin (goat, 1:100, R&D Systems), anti-α₆-integrin (rat, 1:100, Pharmingen), anti-K5 (rabbit, 1:500, Fuchs lab), anti-CD34 (rat, 1:100, Pharmingen), anti-BrdU (rat, 1:500, Abcam), anti-active-caspase 3 (rabbit, 1:500, R&D Systems), anti-TBX1 (rabbit, 1:100, Zymed), and FITC-conjugated (1:100, Jackson) or Alexa594-conjugated (1:1,000, Molecular Probes) secondary antibodies. Nuclei were stained using 4',6-diamidino-2-phenylindole (DAPI). Imaging was performed using either a Zeiss Axioskop equipped with Spot RT (Diagnostic Instruments) or a Zeiss LSM 510 laser-scanning microscope (Carl Zeiss MicroImaging) through a 40× water objective or a 25× objective. RGB images were assembled in Adobe Photoshop CS3, and panels were labelled in Adobe Illustrator CS3.

Isolation of HF-SCs and flow cytometry. Subcutaneous fat was removed from the skins with a scalpel, and the whole skin was placed dermis down on trypsin (Gibco) at 37 °C for 30 min. Single-cell suspensions were obtained by scraping the skin gently. The cells were then filtered with strainers (70 µm, followed by 40 µm). Cell suspensions were incubated with the appropriate antibodies for 30 min on ice. The following antibodies were used: anti-CD34–Alexa647 (1:100, eBioscience) and anti-α₆-integrin–PE (1:100, BD Biosciences). DAPI was used to exclude dead cells. Cell isolations were performed on FACSaria sorters equipped with FACSDiva software (BD Biosciences). FACS analyses were performed using LSRII FACS Analyzers and then analysed with the FlowJo program.

RNA isolation and microarray analyses. RNAs from FACS-purified *Tbx1*-cKO mice and WT HF-SCs 2 days post depilation were provided to the Genomics Core Facility at Memorial Sloan-Kettering Cancer Center for quality control, quantification, reverse transcription, labelling, and hybridization to MOE430A 2.0 microarray chips (Affymetrix). Two entirely independent samples were used for data analyses. Arrays were scanned as per the manufacturer's specifications for the Affymetrix MOE430v2 chip. Images were background-subtracted. Probe sets were identified as differentially expressed when the average fold change was ≥1.8 (*P* < 0.1). Probe sets selected for visualization were log₂ transformed, analysed with hierarchical clustering (Pearson correlation, average linkage) and visualized with heat maps to assist in interpretation.

Noggin intradermal injection. Recombinant mouse noggin (5 µg ml⁻¹, R&D Systems) was injected intradermally, together with beads, into the back skin for 3 days post depilation. The skin was analysed on the third day. BrdU was injected twice in the last 24 h before harvesting the skin.

In vitro BMP4 treatment. For BMP4 treatment experiments, mouse keratinocytes grown in six-well plates and transfected with *Tbx1*-GFP (a *Tbx1* expression vector obtained by cloning a mouse cDNA into the pCMV6-AC-GFP plasmid) were treated with 100 ng ml⁻¹ mouse recombinant BMP4 for the indicated period. After treatment, cells were trypsinized, and GFP⁺ versus GFP⁻ cells were isolated by FACS and separately processed for total RNA extraction and rtPCR analysis.

Immunoblotting. Mouse keratinocytes transduced with scramble or *Smad1* shRNAs by lentiviral infection were selected in 1 µg ml⁻¹ puromycin-containing media for 4 days. Cells were treated with or without BMP4 (100 ng ml⁻¹) for 3 h and lysed directly in SDS sample buffer (50 mM Tris-HCl, pH 6.8, 100 mM DTT, 2% SDS, 0.1% bromophenol blue and 10% glycerol). Gel electrophoresis was performed using 4–12% NuPAGE Bis-Tris gradient gels (Invitrogen), and separated proteins were transferred to PVDF membranes (Millipore). Membranes were blocked for 1 h in Odyssey blocking buffer (LI-COR Biosciences), then

incubated with primary antibodies in blocking buffer overnight at 4 °C. The primary antibodies used were as follows: mouse anti-SMAD1 (1:300, Abcam), rabbit anti-pSMAD1/5/8 (Ser463/465) (1:2,000, Millipore), mouse anti-SMAD2/3 (1:2,000, BD Biosciences), rabbit anti-pSMAD2 (Ser465/467) (1:2,000, Cell Signaling Technology), rabbit anti-SMAD5 (1:500, Abcam) and mouse anti- α -tubulin (1:2,000, Sigma). Secondary antibodies were conjugated to IRDye 680 or IRDye 800 (1:15,000, LI-COR Biosciences).

BMP-reporter experiments. Construction of a lentiviral BMP reporter (pLKO-H2B-CFP-BRE-ZsGreen) is described elsewhere³⁰. Briefly, H2B-CFP works as a transduction marker, which is constitutively transcribed from the *PGK* promoter. As a reporter, ZsGreen is transcribed from a minimal CMV promoter conjugated to the BMP response element (BRE) in the presence of BMP. We transduced the BMP reporter into mouse keratinocytes, FACS-sorted the transduced cells based on H2B-CFP expression and transduced these cells with viral vectors carrying

scramble or *Smad1* shRNAs. After puromycin selection, we tested BMP-reporter activities in the presence or absence of BMP (100 ng ml⁻¹) for 24 h. For visualization and quantification of direct fluorescent signals from the BMP reporter, keratinocytes were seeded onto coverslips and fixed with 4% paraformaldehyde for 10 min at room temperature.

Statistics. To determine the significance between two groups, indicated in the figures by asterisks, comparisons were made using Student's *t*-test, performed by Prism5 software or Microsoft Excel. Box-and-whisker plots are used to describe the entire population without assumptions about the statistical distribution.

31. R Foundation for Statistical Computing. *The R Project for Statistical Computing* (<http://www.r-project.org>) (2011).
32. Beronja, S., Livshits, G., Williams, S. & Fuchs, E. Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nature Med.* **16**, 821–827 (2010).

De novo mutations revealed by whole-exome sequencing are strongly associated with autism

Stephan J. Sanders¹, Michael T. Murtha¹, Abha R. Gupta^{2*}, John D. Murdoch^{1*}, Melanie J. Raubeson^{1*}, A. Jeremy Willsey^{1*}, A. Gulhan Ercan-Sencicek^{1*}, Nicholas M. DiLullo^{1*}, Neelroop N. Parikshak³, Jason L. Stein³, Michael F. Walker¹, Gordon T. Ober¹, Nicole A. Teran¹, Youeun Song¹, Paul El-Fishawy¹, Ryan C. Murtha¹, Murim Choi⁴, John D. Overton⁴, Robert D. Bjornson⁵, Nicholas J. Carriero⁵, Kyle A. Meyer⁶, Kaya Bilguvar⁷, Shrikant M. Mane⁸, Nenad Šestan⁶, Richard P. Lifton⁴, Murat Günel⁷, Kathryn Roeder⁹, Daniel H. Geschwind³, Bernie Devlin¹⁰ & Matthew W. State¹

Multiple studies have confirmed the contribution of rare *de novo* copy number variations to the risk for autism spectrum disorders^{1–3}. But whereas *de novo* single nucleotide variants have been identified in affected individuals⁴, their contribution to risk has yet to be clarified. Specifically, the frequency and distribution of these mutations have not been well characterized in matched unaffected controls, and such data are vital to the interpretation of *de novo* coding mutations observed in probands. Here we show, using whole-exome sequencing of 928 individuals, including 200 phenotypically discordant sibling pairs, that highly disruptive (nonsense and splice-site) *de novo* mutations in brain-expressed genes are associated with autism spectrum disorders and carry large effects. On the basis of mutation rates in unaffected individuals, we demonstrate that multiple independent *de novo* single nucleotide variants in the same gene among unrelated probands reliably identifies risk alleles, providing a clear path forward for gene discovery. Among a total of 279 identified *de novo* coding mutations, there is a single instance in probands, and none in siblings, in which two independent nonsense variants disrupt the same gene, *SCN2A* (sodium channel, voltage-gated, type II, α subunit), a result that is highly unlikely by chance.

We completed whole-exome sequencing in 238 families from the Simons Simplex Collection (SSC), a comprehensively phenotyped autism spectrum disorders (ASD) cohort consisting of pedigrees with two unaffected parents, an affected proband, and, in 200 families, an unaffected sibling⁵. Exome sequences were captured with NimbleGen oligonucleotide libraries, subjected to DNA sequencing on the Illumina platform, and genotype calls were made at targeted bases (Supplementary Information)^{6,7}. On average, 95% of the targeted bases in each individual were assessed by ≥ 8 independent sequence reads; only those bases showing ≥ 20 independent reads in all family members were considered for *de novo* mutation detection. This allowed for analysis of *de novo* events in 83% of all targeted bases and 73% of all exons and splice sites in the RefSeq hg18 database (<http://www.ncbi.nlm.nih.gov/RefSeq/>; Supplementary Table 1; Supplementary Data 1). Given uncertainties regarding the sensitivity of detection of insertion-deletions, case-control comparisons reported here consider only single base substitutions (Supplementary Information). Validation was attempted for all predicted *de novo* single nucleotide variants (SNVs) via Sanger sequencing of all family members, with sequence readers blinded to affected status; 96% were successfully validated. We determined there was no evidence of

systematic bias in variant detection between affected and unaffected siblings through comparisons of silent *de novo*, non-coding *de novo*, and novel transmitted variants (Fig. 1a; Supplementary Figs 1–5; Supplementary Information).

Among 200 quartets (Table 1), 125 non-synonymous *de novo* SNVs were present in probands and 87 in siblings: 15 of these were nonsense (10 in probands; 5 in siblings) and 5 altered a canonical splice site (5 in probands; 0 in siblings). There were 2 instances in which *de novo* SNVs were present in the same gene in two unrelated probands; one of these involved two independent nonsense variants (Table 2). Overall, the total number of non-synonymous *de novo* SNVs was significantly greater in probands compared to their unaffected siblings ($P = 0.01$, two-tailed binomial exact test; Fig. 1a; Table 1) as was the odds ratio (OR) of non-synonymous to silent mutations in probands versus siblings (OR = 1.93; 95% confidence interval (CI), 1.11–3.36; $P = 0.02$, asymptotic test; Table 1). Restricting the analysis to nonsense and splice site mutations in brain-expressed genes resulted in substantially increased estimates of effect size and demonstrated a significant difference in cases versus controls based either on an analysis of mutation burden ($N = 13$ versus 3; $P = 0.02$, two-tailed binomial exact test; Fig. 1a; Table 1) or an evaluation of the odds ratio of nonsense and splice site to silent SNVs (OR = 5.65; 95% CI, 1.44–22.2; $P = 0.01$, asymptotic test; Fig. 1b; Table 1).

To determine whether factors other than diagnosis of ASD could explain our findings, we examined a variety of potential covariates, including parental age, IQ and sex. We found that the rate of *de novo* SNVs indeed increases with paternal age ($P = 0.008$, two-tailed Poisson regression) and that paternal and maternal ages are highly correlated ($P < 0.0001$, two-tailed linear regression). However, although the mean paternal age of probands in our sample was 1.1 years higher than their unaffected siblings, re-analysis accounting for age did not substantively alter any of the significant results reported here (Supplementary Information). Similarly, no significant relationship was observed between the rate of *de novo* SNVs and proband IQ ($P \geq 0.19$, two-tailed linear regression, Supplementary Information) or proband sex ($P \geq 0.12$, two-tailed Poisson regression; Supplementary Fig. 6; Supplementary Information).

Overall, these data demonstrate that non-synonymous *de novo* SNVs, and particularly highly disruptive nonsense and splice-site *de novo* mutations, are associated with ASD. On the basis of the conservative assumption that *de novo* single-base coding mutations observed in siblings confer no autism liability, we estimate that at least 14% of

¹Program on Neurogenetics, Child Study Center, Department of Psychiatry, Department of Genetics, Yale University School of Medicine, 230 South Frontage Road, New Haven, Connecticut 06520, USA.

²Child Study Center, Department of Pediatrics, Yale University School of Medicine, 230 South Frontage Road, New Haven, Connecticut 06520, USA. ³Neurogenetics Program, UCLA, 695 Charles E. Young Dr. South, Los Angeles, California 90095, USA. ⁴Department of Genetics, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ⁵Department of Computer Science, Yale Center for Genome Analysis, Yale University, 51 Prospect Street, New Haven, Connecticut 06511, USA. ⁶Department of Neurobiology, Kavli Institute for Neuroscience, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, USA. ⁷Department of Neurosurgery, Center for Human Genetics and Genomics, Program on Neurogenetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, USA. ⁸Yale Center for Genome Analysis, 300 Heffernan Drive, West Haven, Connecticut 06516, USA. ⁹Department of Statistics, Carnegie Mellon University, 130 DeSoto Street, Pittsburgh, Pennsylvania 15213, USA. ¹⁰Department of Psychiatry and Human Genetics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15213, USA.

*These authors contributed equally to this work.

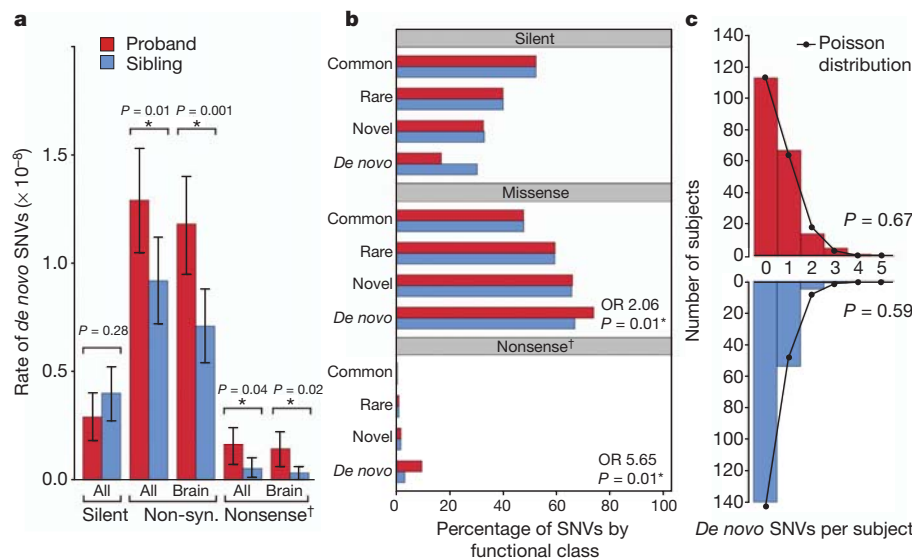


Figure 1 | Enrichment of non-synonymous *de novo* variants in probands relative to sibling controls. **a**, The rate of *de novo* variants is shown for 200 probands (red) and matched unaffected siblings (blue). ‘All’ refers to all RefSeq genes in hg18, ‘Brain’ refers to the subset of genes that are brain-expressed²⁴ and ‘Non-syn’ to non-synonymous SNVs (including missense, nonsense and splice site SNVs). Error bars represent the 95% confidence intervals and *P* values are calculated with a two-tailed binomial exact test. **b**, The proportion of transmitted variants in brain-expressed genes is equal between 200 probands (red) and matched unaffected siblings (blue) for all mutation types and allele frequencies, including common ($\geq 1\%$), rare ($< 1\%$) and novel (single allele in

one of the 400 parents); in contrast, both non-synonymous and nonsense *de novo* variants show significant enrichment in probands compared to unaffected siblings (73.7% versus 66.7%, $P = 0.01$, asymptotic test and 9.5% versus 3.1%, $P = 0.01$ respectively). **c**, The frequency distribution of brain-expressed non-synonymous *de novo* SNVs is shown per sample for probands (red) and siblings (blue). Neither distribution differs from the Poisson distribution (black line), suggesting that multiple *de novo* SNVs within a single individual do not confirm ASD risk. Nonsense[†] represents the combination of nonsense and splice site SNVs.

affected individuals in the SSC carry *de novo* SNV risk events (Supplementary Information). Moreover, among probands and considering brain-expressed genes, an estimated 41% of non-synonymous *de novo* SNVs (95% CI, 21–58%) and 77% of nonsense and splice site

de novo SNVs (95% CI, 33–100%) point to *bona fide* ASD-risk loci (Supplementary Information).

We next set out to evaluate which of the particular *de novo* SNVs identified in our study confer this risk. On the basis of our prior work³,

Table 1 | Distribution of SNVs between probands and siblings

| Category | Total number of SNVs* | | SNVs per subject | | Per base SNV rate (x10 ⁻⁸) | | <i>P</i> † | Odds ratio (95% CI)‡ |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|--|-----------------------|--------------|----------------------|
| | Pro <i>N</i> = 200 | Sib <i>N</i> = 200 | Pro <i>N</i> = 200 | Sib <i>N</i> = 200 | Pro <i>N</i> = 200 | Sib <i>N</i> = 200 | | |
| De novo | | | | | | | | |
| | | | All genes | | | | | |
| All | 154 | 125 § | 0.77 | 0.63 | 1.58 | 1.31 | 0.09 | NA |
| Silent | 29 | 39 | 0.15 | 0.20 | 0.29 | 0.40 | 0.28 | NA |
| All non-synonymous | 125 | 87 | 0.63 | 0.44 | 1.29 | 0.92 | 0.01 | 1.93 (1.11–3.36) |
| Missense | 110 | 82 | 0.55 | 0.41 | 1.13 | 0.86 | 0.05 | 1.80 (1.03–3.16) |
| Nonsense/splice site | 15 | 5 | 0.08 | 0.03 | 0.16 | 0.05 | 0.04 | 4.03 (1.32-12.4) |
| | | | Brain-expressed genes | | | | | |
| All | 137 | 96 | 0.69 | 0.48 | 1.41 | 1.01 | 0.01 | NA |
| Silent | 23 | 30 | 0.12 | 0.15 | 0.24 | 0.31 | 0.41 | NA |
| All non-synonymous | 114 | 67 | 0.57 | 0.34 | 1.18 | 0.71 | 0.001 | 2.22 (1.19–4.13) |
| Missense | 101 | 64 | 0.51 | 0.32 | 1.04 | 0.68 | 0.005 | 2.06 (1.10–3.85) |
| Nonsense/splice site | 13 | 3 | 0.07 | 0.02 | 0.14 | 0.03 | 0.02 | 5.65 (1.44–22.2) |
| Novel transmitted | | | | | | | | |
| | | | All genes | | | | | |
| All | 26,565 | 26,542 | 133 | 133 | 277 | 277 | 0.92 | NA |
| Silent | 8,567 | 8,642 | 43 | 43 | 90 | 91 | 0.57 | NA |
| All non-synonymous | 17,998 | 17,900 | 90 | 90 | 188 | 187 | 0.61 | 1.01 (0.98–1.05) |
| Missense | 17,348 | 17,250 | 87 | 86 | 181 | 180 | 0.60 | 1.01 (0.98–1.05) |
| Nonsense/splice site | 650 | 650 | 3.3 | 3.3 | 7 | 7 | 1.00 | 1.01 (0.90–1.13) |
| | | | Brain-expressed genes | | | | | |
| All | 20,942 | 20,982 | 105 | 105 | 219 | 220 | 0.85 | NA |
| Silent | 6,884 | 6,981 | 34 | 35 | 72 | 74 | 0.42 | NA |
| All non-synonymous | 14,058 | 14,001 | 70 | 70 | 147 | 146 | 0.74 | 1.02 (0.98–1.06) |
| Missense | 13,588 | 13,525 | 68 | 68 | 142 | 141 | 0.71 | 1.02 (0.98–1.06) |
| Nonsense/splice site | 470 | 476 | 2.3 | 2.4 | 5 | 5 | 0.87 | 1.00 (0.88–1.14) |

* An additional 15 *de novo* variants were seen in the probands of 25 trio families; all were missense and 14 were brain-expressed.

† The *P* values compare the number of variants between probands and siblings using a two-tailed binomial exact test (Supplementary Information); *P* values below 0.05 are highlighted in bold.

‡ The odds ratio calculates the proportion of variants in a specific category to silent variants and then compares these ratios in probands versus siblings. NA, not applicable.

§ The sum of silent and non-synonymous variants is 126, however one nonsense and two silent *de novo* variants were identified in *KANK1* in a single sibling, suggesting a single gene conversion event. This event contributed a maximum count of one to any analysis.

Table 2 | Loss of function mutations in probands

| Gene symbol | Gene name | Mutation type |
|-------------|--|---------------|
| ADAM33 | ADAM metallopeptidase domain 33 | Nonsense |
| CSDE1 | cold shock domain containing E1, RNA-binding | Nonsense |
| EPHB2 | EPH receptor B2 | Nonsense |
| FAM8A1 | family with sequence similarity 8, member A1 | Nonsense |
| FREM3 | FRAS1 related extracellular matrix 3 | Nonsense |
| MPHOSPH8 | M-phase phosphoprotein 8 | Nonsense |
| PPM1D | protein phosphatase, Mg ²⁺ /Mn ²⁺ dependent 1D | Nonsense |
| RAB2A | RAB2A, member RAS oncogene family | Nonsense |
| SCN2A | sodium channel, voltage-gated, type II, α subunit | Nonsense |
| SCN2A | sodium channel, voltage-gated, type II, α subunit | Nonsense |
| BTN1A1 | butyrophilin, subfamily 1, member A1 | Splice site |
| FCRL6 | Fc receptor-like 6 | Splice site |
| KATNAL2 | katanin p60 subunit A-like 2 | Splice site |
| NAPRT1 | nicotinate phosphoribosyltransferase domain containing 1 | Splice site |
| RNF38 | ring finger protein 38 | Splice site |
| SCP2 | sterol carrier protein 2 | Frameshift* |
| SHANK2 | SH3 and multiple ankyrin repeat domains 2 | Frameshift* |

*Frameshift *de novo* variants are not included in any of the reported case-control comparisons (Supplementary Information).

we hypothesized that estimating the probability of observing multiple independent *de novo* SNVs in the same gene in unrelated individuals would provide a more powerful statistical approach to identifying ASD-risk genes than the alternative of comparing mutation counts in affected versus unaffected individuals. Consequently, we conducted simulation experiments focusing on *de novo* SNVs in brain-expressed genes, using the empirical data for per-base mutation rates and taking into account the actual distribution of gene sizes and GC content across the genome (Supplementary Information). We calculated probabilities (P) and the false discovery rate (Q) based on a wide range of assumptions regarding the number of genes conferring ASD risk (Supplementary Fig. 7; Fig. 2). On the basis of 150,000 iterations, we determined that under all models, two or more nonsense and/or splice site *de novo* mutations were highly unlikely to occur by chance ($P = 0.008$; $Q = 0.005$; Supplementary Information; Fig. 2a). Importantly, these thresholds were robust both to sample size, and to variation in our estimates of locus heterogeneity. Similarly, in our sample, two or more nonsense or splice site *de novo* mutations remained statistically significant when the simulation was performed using the lower bound of the 95% confidence interval for the estimate of *de novo* mutation rates in probands (Supplementary Fig. 7).

Only a single gene in our cohort, *SCN2A*, met these thresholds ($P = 0.008$; Fig. 2a), with two probands each carrying a nonsense *de novo* SNV (Table 2). This finding is consistent with a wealth of data showing overlap of genetic risks for ASD and seizure⁸. Gain of function mutations in *SCN2A* are associated with a range of epilepsy phenotypes; a nonsense *de novo* mutation has been described in a patient with infantile epileptic encephalopathy and intellectual decline⁹, *de novo* missense mutations with variable electrophysiological effects have been found in cases of intractable epilepsy¹⁰, and transmitted rare missense mutations have been described in families with idiopathic ASD¹¹. Of note, the individuals in the SSC carrying the nonsense *de novo* SNVs have no history of seizure.

We then considered whether alternative approaches described in the recent literature^{4,12}, including identifying multiple *de novo* events in a single individual or predicting the functional consequences of missense mutations, might help identify additional ASD-risk genes. However, we found no differences in the distribution or frequency of multiple *de novo* events within individuals in the case versus the control groups (Fig. 1c). In addition, when we examined patients carrying large *de novo* ASD-risk CNVs, we found a trend towards fewer non-synonymous *de novo* SNVs (Supplementary Fig. 11; Supplementary Information). Consequently, neither finding supported a 'two *de novo* hit' hypothesis. Similarly, we found no evidence that widely used measures of conservation or predictors of protein disruption, such as PolyPhen2¹³, SIFT¹⁴, GERP¹⁵, PhyloP¹⁶ or Grantham Score¹⁷,

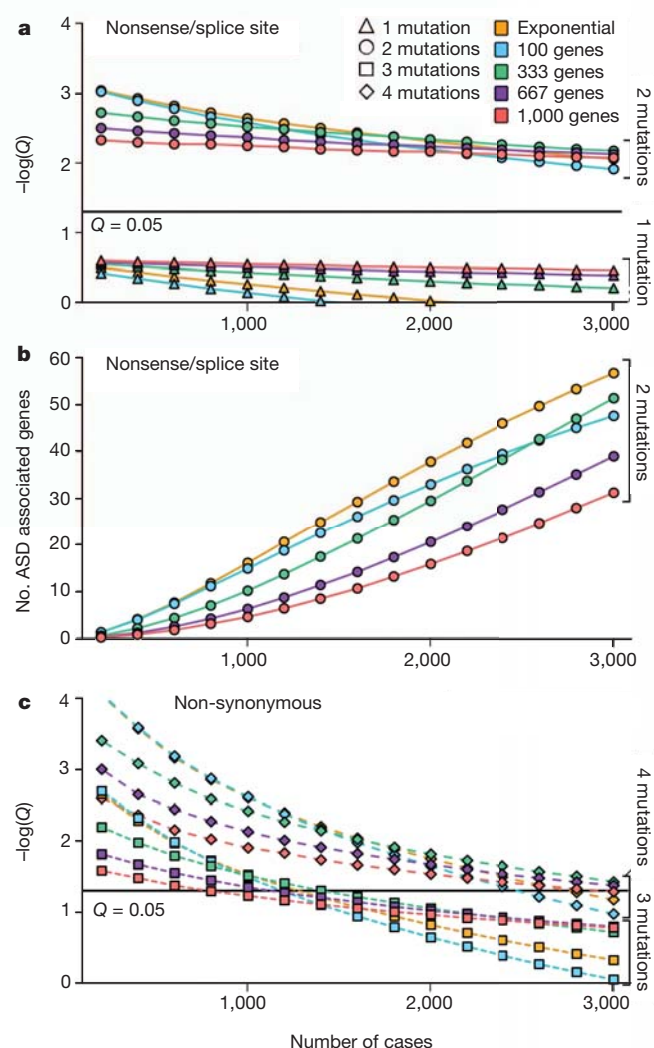


Figure 2 | Identification of multiple *de novo* mutations in the same gene reliably distinguishes risk-associated mutations. **a**, Results of a simulation experiment modelling the likelihood of observing two independent nonsense/splice site *de novo* mutations in the same brain-expressed gene among unrelated probands. We modelled the observed rate of *de novo* brain-expressed mutations in probands and siblings, gene size, GC content and varying degrees of locus heterogeneity, including 100, 333, 667 or 1,000 ASD-contributing genes, as well as using the top 1% of genes derived from a model of exponential distribution of risk (indicated by colour). A total of 150,000 iterations were run. The rate of occurrences of two or more *de novo* variants in non-ASD genes was used to estimate the P -value (Supplementary Fig. 7) while the ratio of occurrences of two or more *de novo* variants in non-ASD genes to similar occurrences in ASD genes was used to estimate the false discovery rate (Q). The identification of two independent nonsense/splice site *de novo* variants in a brain-expressed gene in this sample provides significant evidence for ASD association ($P = 0.008$; $Q = 0.005$) for all models. This observation remained statistically significant when the simulation was repeated using the lower bound of the 95% confidence interval for the estimate of the *de novo* mutation rate in probands (Supplementary Fig. 7). **b**, The simulation described in **a** was used to predict the number of genes that will be found to carry two or more nonsense/splice site *de novo* mutations for a sample of a given size (specified on the x axis). **c**, The simulation was repeated for non-synonymous *de novo* mutations. The identification of three or more independent non-synonymous *de novo* mutations in a brain-expressed gene provides significant evidence for ASD association ($P < 0.05$; $Q < 0.05$) in the sample reported here, however these thresholds are sensitive both to sample size and heterogeneity models.

either alone or in combination differentiated *de novo* non-synonymous SNVs in probands compared to siblings (Supplementary Fig. 9; Supplementary Information). Additionally, among probands, the *de*

de novo SNVs in our study were not significantly over-represented in previously established lists of synaptic genes^{18–20}, genes on chromosome X, autism-implicated genes², intellectual disability genes², genes within ASD-risk associated CNVs³ or *de novo* non-synonymous SNVs identified in schizophrenia probands^{12,21}. Finally we conducted pathway and protein–protein interaction analyses²² for all non-synonymous *de novo* SNVs, all brain-expressed non-synonymous *de novo* SNVs and all nonsense and splice site *de novo* SNVs (Supplementary Fig. 9, 10; Supplementary Information) and did not find a significant enrichment among cases versus controls that survived correction for multiple comparisons, though these studies were of limited power.

These analyses demonstrate that neither the type nor the number of *de novo* mutations observed solely in a single individual provides significant evidence for association with ASD. Moreover, we determined that in the SSC cohort at least three, and most often four or more, brain-expressed non-synonymous *de novo* SNVs in the same gene would be necessary to show a significant association (Fig. 2c; Supplementary Figs 7, 8). Unlike the case of disruptive nonsense and splice site mutations, these simulations were highly sensitive to both sample size and heterogeneity models (Fig. 2c; Supplementary Figs 7, 8; Supplementary Information).

Finally, at the completion of our study, we had the opportunity to combine all *de novo* events in our sample with those identified in an independent whole-exome analysis of non-overlapping Simons Simplex families that focused predominantly on trios²³. From a total of 414 probands, two additional genes were found to carry two highly disruptive mutations each, *KATNAL2* (katanin p60 subunit A-like 2) (our results and ref. 23) and *CHD8* (chromodomain helicase DNA binding protein 8) (ref. 23), thereby showing association with the ASD phenotype.

Overall, our results substantially clarify the genomic architecture of ASD, demonstrate significant association of three genes—*SCN2A*, *KATNAL2* and *CHD8*—and predict that approximately 25–50 additional ASD-risk genes will be identified as sequencing of the 2,648 SSC families is completed (Fig. 2b). Rare non-synonymous *de novo* SNVs are associated with risk, with odds ratios for nonsense and splice-site mutations in the range previously described for large multigenic *de novo* CNVs³. It is important to note that these estimates reflect a mix of risk and neutral mutations in probands. We anticipate that the true effect size for specific SNVs and mutation classes will be further clarified as more data accumulate. From the distribution of large multi-genic *de novo* CNVs in probands versus siblings, we previously estimated the number of ASD-risk loci at 234 (ref. 3). Using the same approach, the current data result in a point estimate of 1,034 genes, however the confidence intervals are large and the distribution of this risk among these loci is unknown (Supplementary Information). What is clear is that our results strongly support a high degree of locus heterogeneity in the SSC cohort, involving hundreds of genes or more. Finally, via examination of mutation rates in well-matched controls, we have determined that the observation of highly disruptive *de novo* SNVs clustering within genes can robustly identify risk-conferring alleles.

The focus on recurrent rare *de novo* mutation described here provided sufficient statistical power to identify associated genes in a relatively small cohort—despite both a high degree of locus heterogeneity and the contribution of intermediate genetic risks. This approach promises to be valuable for future high-throughput sequencing efforts in ASD and other common neuropsychiatric disorders.

METHODS SUMMARY

Sample selection. In total 238 families (928 individuals) were selected from the SSC². Thirteen families (6%) did not pass quality control, leaving 225 families (200 quartets, 25 trios) for analysis (Supplementary Data 1). Of the 200 quartets, 194 (97%) probands had a diagnosis of autism and 6 (3%) were diagnosed with ASD; the median non-verbal IQ was 84.

Exome capture, sequencing and variant prediction. Whole-blood DNA was enriched for exonic sequences through hybridization with a NimbleGen custom array ($N = 210$) or EZExomeV2.0 ($N = 718$). Captured DNA was sequenced using

an Illumina GAIIX ($N = 592$) or HiSeq 2000 ($N = 336$). Short read sequences were aligned to hg18 with BWA⁶, duplicate reads were removed and variants were predicted using SAMtools⁷. Data were normalized within families by only analysing bases with at least 20 unique reads in all family members. *De novo* predictions were made blinded to affected status using experimentally verified thresholds (Supplementary Information). All *de novo* variants were confirmed using Sanger sequencing blinded to affected status.

Gene annotation. Variants were analysed against RefSeq hg18 gene definitions; in genes with multiple isoforms the most severe outcome was chosen. All nonsense and canonical splice site variants were present in all RefSeq isoforms. A variant was listed as altering the splice site only if it disrupted canonical 2-base-pair acceptor (AG) or donor (GT) sites. Brain-expressed genes were identified from expression array analysis across 57 post-mortem brains (age 6 weeks post conception to 82 years) and multiple brain regions; 80% of RefSeq genes were included in this subset²⁴.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 September 2011; accepted 14 February 2012.

Published online 4 April 2012.

- Sebat, J. *et al.* Strong association of *de novo* copy number mutations with autism. *Science* **316**, 445–449 (2007).
- Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368–372 (2010).
- Sanders, S. J. *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70**, 863–885 (2011).
- O’Roak, B. J. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nature Genet.* **43**, 585–589 (2011).
- Fischbach, G. D. & Lord, C. The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron* **68**, 192–195 (2010).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Meisler, M. H., O’Brien, J. E. & Sharkey, L. M. Sodium channel gene family: epilepsy mutations, gene interactions and modifier effects. *J. Physiol. (Lond.)* **588**, 1841–1848 (2010).
- Kamiya, K. *et al.* A nonsense mutation of the sodium channel gene *SCN2A* in a patient with intractable epilepsy and mental decline. *J. Neurosci.* **24**, 2690–2698 (2004).
- Ogiwara, I. *et al.* *De novo* mutations of voltage-gated sodium channel α gene *SCN2A* in intractable epilepsies. *Neurology* **73**, 1046–1053 (2009).
- Weiss, L. A. *et al.* Sodium channels *SCN1A*, *SCN2A* and *SCN3A* in familial autism. *Mol. Psychiatry* **8**, 186–194 (2003).
- Xu, B. *et al.* Exome sequencing supports a *de novo* mutational paradigm for schizophrenia. *Nature Genet.* **43**, 864–868 (2011).
- Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nature Methods* **7**, 248–249 (2010).
- Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols* **4**, 1073–1081 (2009).
- Cooper, G. M. *et al.* Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. *Nature Methods* **7**, 250–251 (2010).
- Cooper, G. M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* **15**, 901–913 (2005).
- Grantham, R. Amino acid difference formula to help explain protein evolution. *Science* **185**, 862–864 (1974).
- Abul-Husn, N. S. *et al.* Systems approach to explore components and interactions in the presynapse. *Proteomics* **9**, 3303–3315 (2009).
- Bayés, A. *et al.* Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nature Neurosci.* **14**, 19–21 (2011).
- Collins, M. O. *et al.* Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. *J. Neurochem.* **97** (suppl. 1), 16–23 (2006).
- Girard, S. L. *et al.* Increased exonic *de novo* mutation rate in individuals with schizophrenia. *Nature Genet.* **43**, 860–863 (2011).
- Rossin, E. J. *et al.* Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet.* **7**, e1001273 (2011).
- O’Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* <http://dx.doi.org/10.1038/nature10989> (this issue).
- Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to all of the families participating in the Simons Foundation Autism Research Initiative (SFARI) Simplex Collection (SSC). This work was supported by a grant from the Simons Foundation. R.P.L. is an Investigator of the

Howard Hughes Medical Institute. We thank the SSC principal investigators A. L. Beaudet, R. Bernier, J. Constantino, E. H. Cook Jr, E. Fombonne, D. Geschwind, D. E. Grice, A. Klin, D. H. Ledbetter, C. Lord, C. L. Martin, D. M. Martin, R. Maxim, J. Miles, O. Ousley, B. Peterson, J. Piggot, C. Saulnier, M. W. State, W. Stone, J. S. Sutcliffe, C. A. Walsh and E. Wijsman and the coordinators and staff at the SSC sites for the recruitment and comprehensive assessment of simplex families; the SFARI staff, in particular M. Benedetti, for facilitating access to the SSC; Prometheus Research for phenotypic data management and Prometheus Research and the Rutgers University Cell and DNA repository for accessing biomaterials; the Yale Center of Genomic Analysis, in particular M. Mahajan, S. Umlauf, I. Tikhonova and A. Lopez, for generating sequencing data; T. Brooks-Boone, N. Wright-Davis and M. Wojciechowski for their help in administering the project at Yale; I. Hart for support; G. D. Fischbach, A. Packer, J. Spiro, M. Benedetti and M. Carlson for their suggestions throughout; and B. Neale and M. Daly for discussions regarding *de novo* variation. We also acknowledge T. Lehner and the Autism Sequencing Consortium for providing an opportunity for pre-publication data exchange among the participating groups.

Author Contributions S.J.S., M.T.M., R.P.L., M.G., D.H.G. and M.W.S. designed the study; M.T.M., A.R.G., J.M., M.R., A.G.E.-S., N.M.D., S.M., M.W., G.O., Y.S., P.E., R.M. and J.O. designed and performed high-throughput sequencing experiments and variant confirmations; S.J.S., M.C., K.B., R.B. and N.C. designed the exome-analysis bioinformatics pipeline; S.J.S., A.J.W., N.N.P., J.L.S., N.T., K.A.M., N.S., K.R., D.H.G., B.D. and M.W.S. analysed the data; S.J.S., A.J.W., K.R., B.D. and M.W.S. wrote the paper; J.M., M.R., A.J.W., A.R.G., A.G.E.-S. and N.M.D. contributed equally to the study. All authors discussed the results and contributed to editing the manuscript.

Author Information Sequence data from this study is available through the NCBI Sequence Read Archive (accession number SRP010920.1). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.W.S. (matthew.state@yale.edu), B.D. (devlinbj@upmc.edu) or D.H.G. (dhg@mednet.ucla.edu).

METHODS

Sample selection. In total 238 families (928 individuals) were selected from the SSC on the basis of: male probands with autism, low non-verbal IQ (NVIQ), and discordant Social Responsiveness Scale (SRS) with sibling and parents ($N = 40$); female probands ($N = 46$); multiple unaffected siblings ($N = 28$); probands with known multigenic CNVs ($N = 15$); and random selection ($N = 109$). Thirteen families (6%) did not pass quality control (Supplementary Information) leaving 225 families (200 quartets, 25 trios) for analysis (Supplementary Data 1). Of the 200 quartets, 194 (97%) probands had a diagnosis of autism and 6 (3%) were diagnosed with ASD; the median NVIQ was 84. Three of these quartets have previously been reported as trios⁴; there is no overlap between the current sample and those presented in the companion article²³.

Exome capture, sequencing and variant prediction. Whole-blood DNA was enriched for exonic sequences (exome capture) through hybridization with a NimbleGen custom array ($N = 210$) or EZExomeV2.0 ($N = 718$). The captured DNA was sequenced using an Illumina GAIIX ($N = 592$) or HiSeq 2000 ($N = 336$). Short read sequences were aligned to hg18 with BWA⁶, duplicate reads were removed and variants were predicted using SAMtools⁷. The data were normalized across each family by only analysing bases with at least 20 unique reads in all family members (Supplementary Information). *De novo* predictions were made blinded to affected status using experimentally verified thresholds (Supplementary Information). All *de novo* variants were confirmed using Sanger sequencing blinded to affected status.

Variant frequency. The allele frequency of a given variant in the offspring was determined by comparison with dbSNPv132 and 1,637 whole-exome controls including 400 parents. Variants were classified as: 'novel', if only a single allele was present in a parent and none were seen in dbSNP or the other control exomes; 'rare', if they did not meet the criteria for novel and were present in <1% of controls; and 'common', if they were present in ≥1% of controls.

Gene annotation. Variants were analysed against the RefSeq hg18 gene definitions, a list that includes 18,933 genes. Where multiple isoforms gave varying results the most severe outcome was chosen. All nonsense and canonical splice site variants were checked manually and were present in all RefSeq isoforms. A variant was listed as altering the splice site only if it disrupted canonical 2-base-pair acceptor (AG) or donor (GT) sites.

Brain-expressed genes. A list of brain-expressed genes was obtained from expression array analysis across 57 post-mortem brains (age 6 weeks post conception to 82 years) and multiple brain regions²⁴. Using these data, 14,363 (80%) of genes were classified as brain-expressed (Supplementary Information).

Rate of *de novo* SNVs. To allow an accurate comparison between the *de novo* burden in probands and siblings, the number of *de novo* SNVs found in each sample was divided by the number of bases analysed (that is, bases with ≥20 unique reads in all family members) to calculate a per-base rate of *de novo* SNVs. Rates are given in Table 1.

Simulation model. The likelihood of observing multiple independent *de novo* events of a given type for a given sample size in an ASD risk-conferring gene was modelled using gene size and GC content (derived from the full set of brain-expressed RefSeq genes) and the observed rate of brain-expressed *de novo* variants in probands and siblings. These values were then used to evaluate the number of genes contributing to ASD showing two or more variants of the specified type (Fig. 2); comparing this to the number of genes with similar events not carrying ASD risk gave the likelihood of the specified pattern demonstrating association with ASD. The simulation was run through 150,000 iterations across a range of samples sizes and multiple models of locus heterogeneity (Supplementary Information).

Severity scores. Severity scores were calculated for missense variants using web-based interfaces for PolyPhen2¹³, SIFT¹⁴ and GERP¹⁵, using the default settings (Supplementary Information). PhyloP¹⁶ and Grantham Score¹⁷ were determined using an in-house annotated script. For nonsense/splice site variants the maximum score was assigned for Grantham, SIFT and PolyPhen2; for GERP and PhyloP, every possible coding base for the specific protein was scored and the highest value selected.

Pathway analysis. The list of brain-expressed genes with non-synonymous *de novo* SNVs was submitted to KEGG using the complete set of 14,363 brain-expressed genes as the background to prevent bias. For IPA the analysis was based on human nervous system pathways only, again to prevent bias. Otherwise default settings were used for both tools.

Protein-protein interactions. Genes with brain-expressed non-synonymous *de novo* variants in probands were submitted to the Disease Association Protein-protein Link Evaluator (DAPPLE)²² using the default settings.

Comparing *de novo* SNV counts to gene lists. To assess whether non-synonymous *de novo* SNVs were enriched in particular gene sets, the chance of seeing a *de novo* variant in each gene on a given list was estimated based on the size and GC content of the gene. The observed number of *de novo* events was then assessed using the binomial distribution probability based on the total number of non-synonymous *de novo* variants in probands and the sum of probabilities for *de novo* events within these genes.

An inverse relationship to germline transcription defines centromeric chromatin in *C. elegans*

Reto Gassmann^{1*}, Andreas Rechtsteiner^{2*}, Karen W. Yuen^{1†*}, Andrew Muroyama¹, Thea Egelhofer², Laura Gaydos², Francie Barron^{1†}, Paul Maddox^{1†}, Anthony Essex^{1†}, Joost Monen^{1†}, Sevinc Ercan³, Jason D. Lieb³, Karen Oegema¹, Susan Strome² & Arshad Desai¹

Centromeres are chromosomal loci that direct segregation of the genome during cell division. The histone H3 variant CENP-A (also known as CenH3) defines centromeres in monocentric organisms, which confine centromere activity to a discrete chromosomal region, and holocentric organisms, which distribute centromere activity along the chromosome length^{1–3}. Because the highly repetitive DNA found at most centromeres is neither necessary nor sufficient for centromere function, stable inheritance of CENP-A nucleosomal chromatin is postulated to propagate centromere identity epigenetically⁴. Here, we show that in the holocentric nematode *Caenorhabditis elegans* pre-existing CENP-A nucleosomes are not necessary to guide recruitment of new CENP-A nucleosomes. This is indicated by lack of CENP-A transmission by sperm during fertilization and by removal and subsequent reloading of CENP-A during oogenic meiotic prophase. Genome-wide mapping of CENP-A location in embryos and quantification of CENP-A molecules in nuclei revealed that CENP-A is incorporated at low density in domains that cumulatively encompass half the genome. Embryonic CENP-A domains are established in a pattern inverse to regions that are transcribed in the germline and early embryo, and ectopic transcription of genes in a mutant germline altered the pattern of CENP-A incorporation in embryos. Furthermore, regions transcribed in the germline but not embryos fail to incorporate CENP-A throughout embryogenesis. We propose that germline transcription defines genomic regions that exclude CENP-A incorporation in progeny, and that zygotic transcription during early embryogenesis remodels and reinforces this basal pattern. These findings link centromere identity to transcription and shed light on the evolutionary plasticity of centromeres.

To characterize CENP-A localization dynamics in *C. elegans* (CeCENP-A; Supplementary Figs 1 and 2), we generated a strain in which the only source of CENP-A is a single copy green fluorescent protein (GFP)-conjugated transgene encoding GFP-CeCENP-A (Supplementary Fig. 1b). Imaging in adult hermaphrodites revealed that, in the maternal germline, CeCENP-A is removed from chromosomes as they enter the pachytene stage of meiotic prophase, and is reloaded when nuclei progress into diplotene (Fig. 1a). CeCENP-A was not detected in the nuclei of mature sperm (Supplementary Fig. 2a, b), and quantitative immunoblotting indicated that sperm have fewer than the detection limit of 300 CeCENP-A molecules (Fig. 1b and Supplementary Fig. 2c, d). To test for sperm-derived CeCENP-A in embryos, we fertilized CeCENP-A-depleted oocytes with wild-type sperm. In control embryos, CeCENP-A localized to both sperm and oocyte chromatin during chromosome condensation, pronuclear migration and mitosis

(Fig. 1c). As reported in other systems⁵, this recruitment was independent of DNA replication (Supplementary Fig. 3). After fertilization of CeCENP-A-depleted oocytes with wild-type sperm, no CeCENP-A signal was detected on sperm or oocyte chromatin throughout the cell cycle (Fig. 1c). Thus, sperm chromatin does not retain CeCENP-A to propagate centromere identity through fertilization in *C. elegans*.

Pulse-chase experiments in human cells have suggested that stable inheritance of CENP-A on chromatin propagates centromere identity through cell division^{6,7}. To test whether CeCENP-A is stably inherited on chromatin in embryos, we photobleached one set of GFP-CeCENP-A-labelled chromatids after separation from their sisters at anaphase onset in the one-cell embryo. GFP-CeCENP-A signals were then compared in the next round of division between cells inheriting bleached or unbleached chromatid sets. Stable inheritance of CeCENP-A on chromatin predicts that 50% of CeCENP-A is old and the other 50% is new, resulting in a bleached/unbleached ratio of 0.5. In contrast, if CeCENP-A is not stably inherited on chromatin between the two rounds of division, the bleached/unbleached ratio should be 1.0, which is close to the observed value (Fig. 1d–f and Supplementary Fig. 4). Thus, despite the short division time (only ~15 min between consecutive metaphases), CeCENP-A is nearly completely turned over on chromatin during embryonic cell divisions.

The above results indicate that pre-existing CeCENP-A nucleosomes may not be the cue that targets new CeCENP-A nucleosomes. To define the unknown guiding cue(s), we analysed the genome-wide distribution of CeCENP-A in embryos using chromatin immunoprecipitation with a CeCENP-A-specific antibody⁸ followed by hybridization to a tiling microarray (ChIP-chip). As CENP-A chromatin is characterized by highly repetitive DNA in most higher eukaryotes, this offered the opportunity to define the distribution of CENP-A in an organism naturally lacking large stretches of repeats. Our ChIP-chip analysis revealed regions of CeCENP-A enrichment along the entire length of chromosomes, as predicted for holocentric chromosome architecture (Fig. 2a and Supplementary Fig. 5a–c). The genome-wide distribution of the conserved CENP-A-specific loading factor KNL-2 was indistinguishable from that of CeCENP-A, indicating that the CeCENP-A distribution reflects specific incorporation (Fig. 2a, b). A sliding-window-based domain definition algorithm revealed that CeCENP-A domains vary considerably in size (median 10–12 kilobase), cover both genic and intergenic regions, are distributed evenly throughout the genome, and do not correlate with repeat density (Fig. 2c–e and Supplementary Fig. 5d–f). Although nearly half the genome is occupied by CeCENP-A domains (Fig. 2d), quantification of CeCENP-A molecules in purified embryonic nuclei showed that there is only enough CeCENP-A to occupy at most 4% of the genome

¹Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine University of California San Diego, La Jolla, California 92037, USA. ²Department of Molecular Cell and Developmental Biology, University of California Santa Cruz, Santa Cruz, California 95064, USA. ³Department of Biology, Carolina Center for Genome Sciences and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [†]Present addresses: School of Biological Sciences, the University of Hong Kong, Pokfulam, Hong Kong (K.W.Y.); Department of Craniofacial Biology, University of Colorado, Health Sciences Center, Aurora, Colorado 80045, USA (F.B.); Institute for Research in Immunology and Cancer, Department of Pathology and Cell Biology, University of Montreal, Montreal, Quebec H3C 3J7, Canada (P.M.); The Salk Institute for Biological Studies, San Diego, California 92186, USA (A.E.); Department of Neurobiology and Behavior, Cornell University, Ithaca, New York 14853, USA (J.M.).

*These authors contributed equally to this work.

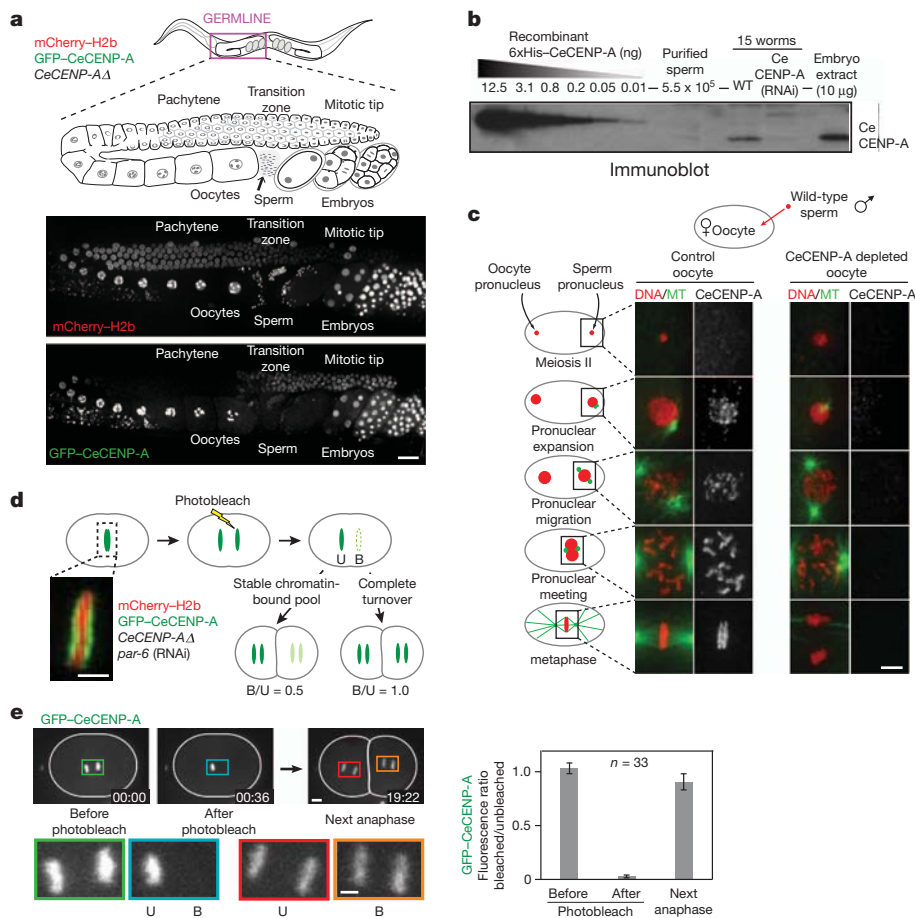


Figure 1 | CeCENP-A dynamics in meiotic prophase, at fertilization and across embryonic divisions. **a**, Gonad region of an adult hermaphrodite co-expressing GFP-CeCENP-A and mCherry-histone H2b in a *CeCENP-A* Δ (also known as *hcp-3(ok1892)*; see also Supplementary Figs 1 and 2) background (see also Supplementary Figs 1 and 2). Scale bar, 20 μ m. **b**, Quantitative immunoblot showing that sperm lack a significant pool of CeCENP-A (see also Supplementary Fig. 2c, d). **c**, Fertilized one-cell control or CeCENP-A-depleted embryos at different stages of the first mitotic division were immunostained for CeCENP-A and α -tubulin (MT). Wild-type (N2) males were mated to *fem-1* mutant worms to ensure all embryos were cross-progeny. Scale bar, 5 μ m. **d**, Schematic of photobleaching experiment to assay CeCENP-A inheritance across early embryonic divisions. *par-6* RNA interference (RNAi) abolishes developmental asynchrony in the two-cell embryo. Unbleached (U) and bleached (B) chromatid sets are indicated. Scale bar, 2 μ m. **e**, Representative images and quantification of the photobleaching experiment. Higher magnification views highlight bleached and unbleached chromatid sets. Error bars are 95% confidence intervals for the means. Scale bars, 5 μ m.

(Fig. 2e and Supplementary Fig. 6a–c). Therefore, the domains enriched for CeCENP-A identified by ChIP-chip must be comprised primarily of H3 nucleosomes (Fig. 2f). Consistent with this, histone H3 ChIP-chip analysis does not show depletion in regions enriched for CeCENP-A⁹ (Supplementary Fig. 6d, e). Thus, the holocentric architecture of *C. elegans* chromosomes arises from reproducible definition of domains that are permissive for low-density CeCENP-A incorporation.

The abundance of genomic regions permissive for CeCENP-A incorporation makes it unlikely that they are defined by a specific DNA sequence. Instead, a correlation emerged with transcriptional status: genes transcribed in embryos were refractory to CeCENP-A incorporation, whereas genes that are silent in embryos (but transcribed in post-embryonic tissues) were permissive for CeCENP-A incorporation (Fig. 3a). ChIP-chip analysis of RNA polymerase II (Pol II) revealed an inverse correlation with CeCENP-A that extended genome-wide (Fig. 3b, c and Supplementary Fig. 7).

The inverse correlation between transcription and CeCENP-A incorporation was puzzling, given that there is no significant RNA Pol II-dependent transcription during the first two rounds of embryonic division, and transcriptional activity remains low until the 30-cell stage^{10–12}. In addition, inhibition of transcription using α -amanitin did not cause defects in chromosome segregation in early embryos (Supplementary Fig. 8). We analysed the CeCENP-A and RNA Pol II distribution in four populations of embryos that formed a developmentally timed series from very early (73% of embryos with ≤ 8 nuclei) to old (67% of embryos with > 200 nuclei) (Supplementary Fig. 9a, b). The CeCENP-A distribution remained constant across this series, despite the activation or repression of genes (Supplementary Figs 9b–d and 10a, b). Thus, CeCENP-A incorporation in embryos may not be dictated simply by active transcription.

To better assess the relationship between transcription and CeCENP-A incorporation, we analysed CeCENP-A and RNA Pol II

enrichment in different gene classes defined in previous work by their expression profiles¹³. The inverse correlation between CeCENP-A and RNA Pol II held true for four of five gene classes (Fig. 3d). However, the ‘germline-only’ class, which is comprised of genes transcribed in the maternal germline but not in embryos, did not show an inverse correlation. Instead, both CeCENP-A and embryonic RNA Pol II levels were low (Fig. 3e), which was confirmed by individual analysis of well-characterized genes transcribed exclusively in the germline (Fig. 3f). In addition, ‘germline-only’ genes failed to incorporate CeCENP-A throughout embryogenesis, despite the persistent absence of RNA Pol II (Supplementary Fig. 10a, b). Germline-only genes are enriched for histone H3 lysine 36 methylation (H3K36me), indicating that the absence of CeCENP-A and RNA Pol II signals is not a false negative¹³ (Fig. 3d–f). Thus, genes transcribed in the maternal germline are refractory to CeCENP-A incorporation, even if they are transcriptionally silent in embryos. This result, together with the paucity of transcription in early embryos and the consistent CeCENP-A distribution throughout embryogenesis, indicates that transcriptional activity in the maternal germline may render genomic regions refractory to CeCENP-A incorporation. Activation of transcription during early embryogenesis probably remodels and reinforces this basal CeCENP-A pattern. In support of this, genes transcribed in early embryos, but lacking signatures of germline transcription, also show low CeCENP-A occupancy (Supplementary Fig. 10c).

If germline transcriptional activity influences CeCENP-A incorporation in the progeny, changes in germline transcription should alter the CeCENP-A distribution in embryos. An unexpected observation in embryos derived from a null mutant of *met-1*, which encodes one of two *C. elegans* H3K36 methyltransferases¹⁴, enabled us to test this prediction. The *met-1* mutant is viable and fertile, indicating that transcription is not globally misregulated in this mutant. Consistent with this, the genome-wide distributions of H3K36me3, RNA Pol II

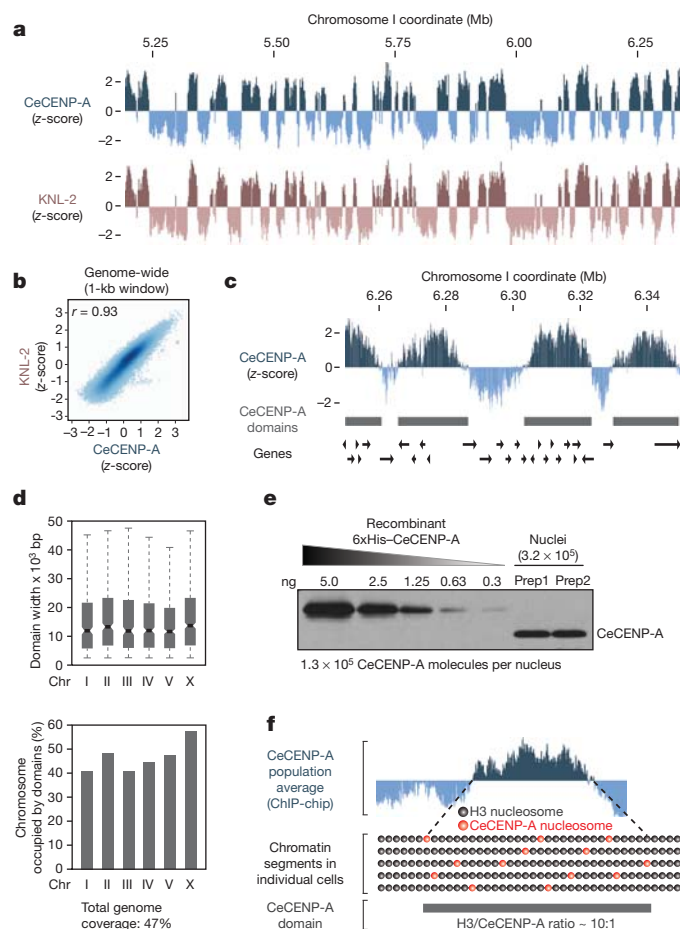


Figure 2 | Genome-wide mapping of CeCENP-A-enriched chromatin.

a, Regions enriched for CeCENP-A and its loading factor KNL-2 in a representative portion of chromosome I. For each track, the average z-score probe signal of two independent biological replicates is plotted. **b**, Genome-wide correlation plot of CeCENP-A and KNL-2 occupancy. The correlation coefficient (r) is in the upper left corner. **c**, Regions enriched for CeCENP-A with the positions of annotated genes. CeCENP-A domains were defined by a sliding window algorithm. **d**, Features of CeCENP-A domains for individual chromosomes. Boxplots: boxes indicate 25th to 75th percentile, whiskers 2.5th to 97.5th percentile. Wedges around the medians indicate 95% confidence intervals for the medians (see also Supplementary Fig. 5d–f). **e**, Two independent nuclei preparations (Prep) from early embryos (<100 nuclei) were blotted alongside a purified CeCENP-A standard (see also Supplementary Figs 2c, d and 6a–c). **f**, Hypothetical model for CeCENP-A permissive domain.

and CeCENP-A were similar in *met-1* and wild-type embryos (Fig. 4a; genome-wide correlation coefficients of 0.88, 0.9 and 0.9, respectively). However, we observed rare regions of ectopic H3K36me3 enrichment in the *met-1* mutant (Fig. 4 and Supplementary Figures 11 and 12). Out of 132 regions >5 kb in size that acquired ectopic H3K36me3 signal in *met-1* mutant embryos, 75 did not show significant RNA Pol II occupancy in embryos, suggesting that these regions are mis-transcribed in mutant germlines but not in embryos. CeCENP-A was depleted from these regions in embryos (Fig. 4a, b and Supplementary Figs 11 and 12a). To test if acquisition of H3K36me3 and loss of CeCENP-A in *met-1* mutant embryos is associated with ectopic germline gene transcription, we hand-dissected germlines from adult wild-type and *met-1* mutant worms and measured messenger RNA levels by quantitative PCR for nine genes in ectopic H3K36me3 regions. As controls, we used eight genes located in regions that did not show a change in H3K36me3 signal. Genes in regions with ectopic H3K36me3 signal indeed showed significantly elevated RNA levels in *met-1* mutant germlines compared to wild type (Fig. 4b and Supplementary Fig. 12b). Thus, the data obtained with the *met-1* mutant indicate

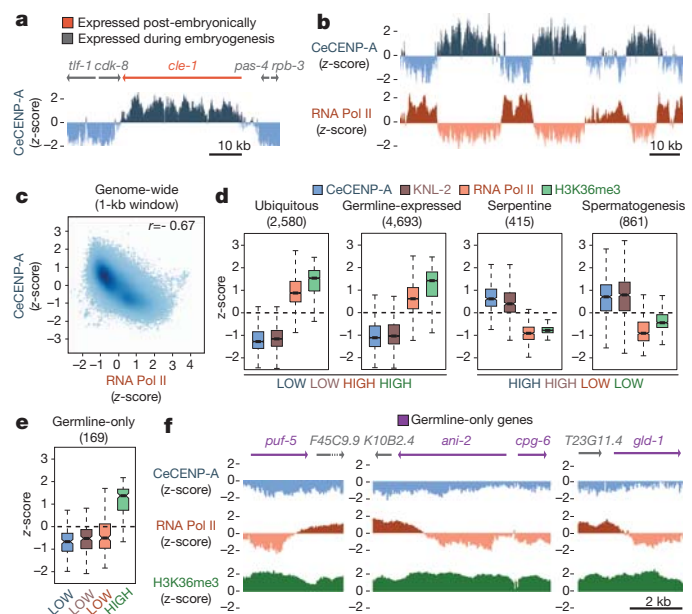


Figure 3 | Relationship between CeCENP-A and gene expression.

a, Chromosomal region containing the *cle-1* gene, which is expressed in neurons, and flanking genes that are expressed during embryogenesis. **b**, Genome browser view showing inverse correlation between CeCENP-A and RNA Pol II occupancy. **c**, Genome-wide correlation plot of CeCENP-A and RNA Pol II occupancy. The correlation coefficient (r) is in the upper right corner. **d**, CeCENP-A, KNL-2, RNA Pol II and H3K36me3 occupancy for various gene sets defined on the basis of expression data. The number of genes in each set is shown in parentheses. Boxplots (as in Fig. 2d) show the range of z-scores averaged over gene bodies. **e**, CeCENP-A, KNL-2, RNA Pol II and H3K36me3 occupancy for the germline-only gene set. **f**, Genome browser views of CeCENP-A, RNA Pol II, and H3K36me3 occupancy on germline-only genes, flanked by genes expressed in embryos.

that ectopic transcription in the germline converts regions from permissive to non-permissive for CeCENP-A incorporation in the embryo progeny.

Whereas the genome-wide inverse relationship to embryonic transcription suggests the simple model that CeCENP-A deposition is random and antagonized by active transcription, such a model fails to explain restriction of *de-novo*-deposited CeCENP-A on transcriptionally silent sperm chromatin (Fig. 1d), the low CeCENP-A occupancy on the 169 ‘germline-only’ genes throughout embryogenesis, and the results from the *met-1* mutant analysis (Supplementary Discussion). Thus, we favour the model that transcription in the germline makes regions non-permissive for CeCENP-A incorporation in the progeny (Supplementary Fig. 12c), and the onset of transcriptional activity in embryos reinforces and remodels this pattern.

In *C. elegans*, H3K36 methylation and the Argonaute CSR-1, which binds short 22G-RNAs (named for their 5' guanosine residue and 22-nucleotide length) derived from germline transcripts, are candidate mechanisms for transmitting memory of germline transcription to early embryos^{13,15}. Both H3K36 methylation (Fig. 3d–f and Supplementary Fig. 7c) and CSR-1 22G-RNA targets (Supplementary Fig. 13) are inversely correlated with CeCENP-A occupancy, and inhibition of CSR-1 and its co-factors leads to early embryonic chromosome segregation defects^{15,16}.

The results here demonstrate that cues unrelated to pre-existing CENP-A nucleosomes can dictate the incorporation of new CENP-A nucleosomes, challenging the view that centromeres are patterned by stable inheritance of CENP-A domains with the key mark for centromere identity being CENP-A itself. Discontinuity of chromatin-localized CENP-A in the germline, similar to the one we describe here for *C. elegans*, has recently been proposed to also occur in plants¹⁷, suggesting that removal and reloading of CENP-A during every

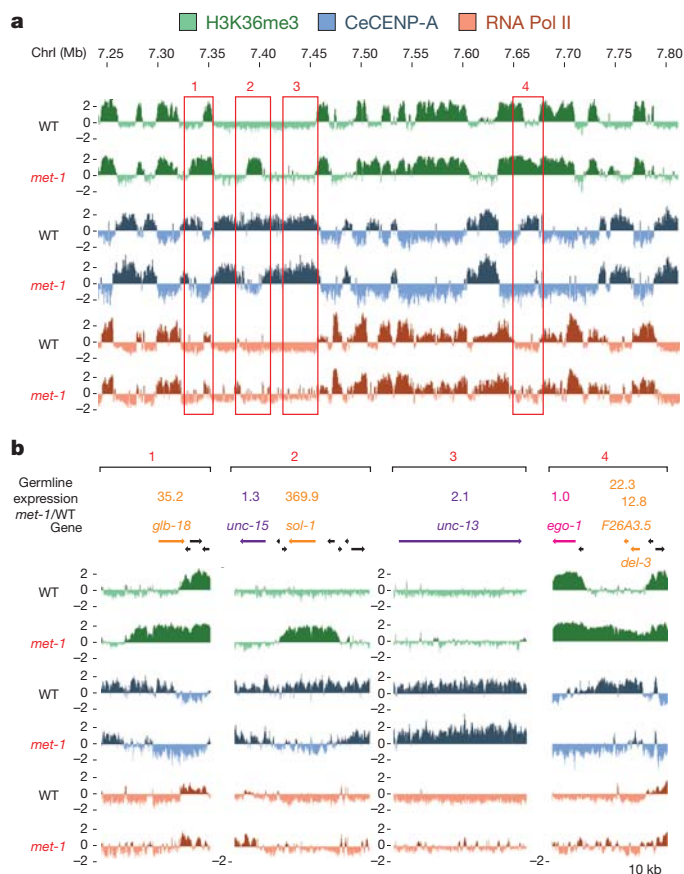


Figure 4 | Germline expression controls CeCENP-A occupancy in the progeny embryos. **a**, Portion of chromosome I featuring specific regions with ectopic H3K36me3 signal in the *met-1* mutant (see also Supplementary Figs 11 and 12). **b**, Screen shots of the regions boxed in (**a**). Real-time quantitative reverse transcription PCR was performed on hand-dissected wild-type and *met-1* mutant gonads. Mean *met-1*:wild-type expression ratio (four independent biological replicates each) is listed above genes (see table in Supplementary Fig. 12b for all genes analysed).

generation may be more common than is currently appreciated. In addition, the processes of neocentromerization and centromere repositioning, which occur with appreciable frequency in humans and are observed frequently during evolution^{18–20}, may be guided by cues that are linked to transcription.

METHODS SUMMARY

For ChIP-chip, chitinase-treated embryos were fixed with 1% formaldehyde in PBS for 10 min, suspended in ChIP buffer (50 mM HEPES-KOH pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.5% NP-40, 0.1% deoxycholate, 1% sarkosyl), and sonicated with a Branson sonifier microtip. Antibody (5 µg) pre-bound to 50 µl Dynabeads (Dyna Beads) were incubated for 4 h with embryo extract (3 mg total protein). Beads were washed and eluted, and the purified DNA amplified as described¹³ before hybridization at the Roche Nimblegen Service Laboratory (2.1M probe tiling arrays with 50-bp probes; WormBase version WS170). Genome-wide scatter plots and Pearson correlations were obtained using log₂ z-scores after median smoothing over 1-kb windows. GFP–CeCENP-A images were acquired with a CSU10 spinning disk confocal head (Yokogawa) and a CCD camera (iXon DV887; Andor Technology) mounted on a Nikon TE2000-E inverted microscope equipped with a solid state laser combiner (ALC) 491 nm and 561 nm lines. For quantification of CeCENP-A in nuclei, embryos were treated with chitinase and lysed by douncing in nuclei buffer (10 mM Tris-HCl pH 8, 80 mM KCl, 2 mM K-EDTA, 0.75 mM spermidine, 0.3 mM spermine, 0.1% digitonin). Nuclei were separated from debris by low-speed centrifugation steps. For germline expression analysis, total RNA from 50–100 dissected gonads was isolated using TRIzol (Invitrogen), and complementary DNA was synthesized with Superscript III reverse transcriptase (Invitrogen).

Quantitative real-time PCR was performed with iQ SYBR Green Supermix (Bio-Rad) in the iQ5 cyclor (Bio-Rad) using standard protocols.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 October 2011; accepted 16 February 2012.

Published online 8 April 2012.

- Malik, H. S. & Henikoff, S. Major evolutionary transitions in centromere complexity. *Cell* **138**, 1067–1082 (2009).
- Allshire, R. C. & Karpen, G. H. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nature Rev. Genet.* **9**, 923–937 (2008).
- Choo, K. *The Centromere* (Oxford Univ. Press, 1997).
- Sullivan, K. F. A solid foundation: functional specialization of centromeric chromatin. *Curr. Opin. Genet. Dev.* **11**, 182–188 (2001).
- Shelby, R. D., Monier, K. & Sullivan, K. F. Chromatin assembly at kinetochores is uncoupled from DNA replication. *J. Cell Biol.* **151**, 1113–1118 (2000).
- Jansen, L. E. T., Black, B. E., Foltz, D. R. & Cleveland, D. W. Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* **176**, 795–805 (2007).
- Schuh, M., Lehner, C. F. & Heidmann, S. Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr. Biol.* **17**, 237–243 (2007).
- Monen, J., Maddox, P. S., Hyndman, F., Oegema, K. & Desai, A. Differential role of CENP-A in the segregation of holocentric *C. elegans* chromosomes during meiosis and mitosis. *Nature Cell Biol.* **7**, 1248–1255 (2005).
- Liu, T. *et al.* Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Res.* **21**, 227–236 (2011).
- Seydoux, G. & Dunn, M. A. Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. *Development* **124**, 2191–2201 (1997).
- Baugh, L. R., Hill, A. A., Slonim, D. K., Brown, E. L. & Hunter, C. P. Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome. *Development* **130**, 889–900 (2003).
- Edgar, L. G., Wolf, N. & Wood, W. B. Early transcription in *Caenorhabditis elegans* embryos. *Development* **120**, 443–451 (1994).
- Reichsteiner, A. *et al.* The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. *PLoS Genet.* **6**, e1001091 (2010).
- Andersen, E. C. & Horvitz, H. R. Two *C. elegans* histone methyltransferases repress *lin-3* EGF transcription to inhibit vulval development. *Development* **134**, 2991–2999 (2007).
- Claycomb, J. M. *et al.* The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell* **139**, 123–134 (2009).
- van Wolfswinkel, J. C. *et al.* CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell* **139**, 135–148 (2009).
- Ravi, M. *et al.* Meiosis-specific loading of the centromere-specific histone CENH3 in *Arabidopsis thaliana*. *PLoS Genet.* **7**, e1002121 (2011).
- Lomiento, M., Jiang, Z., D'Addabbo, P., Eichler, E. E. & Rocchi, M. Evolutionary-new centromeres preferentially emerge within gene deserts. *Genome Biol.* **9**, R173 (2008).
- Piras, F. M. *et al.* Uncoupling of satellite DNA and centromeric function in the genus *Equus*. *PLoS Genet.* **6**, e1000845 (2010).
- Warburton, P. E. Chromosomal dynamics of human neocentromere formation. *Chromosome Res.* **12**, 617–626 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Ward for the purified sperm sample, M. Gupta for help with analysis, J. Ahinger for advice on fixation and ChIP procedures, and A. Dernburg and other members of the Lieb modENCODE group for helpful discussions. This work was supported by a modENCODE grant (U01 HG004270), and by grants from NIH to A.D. (GM074215 and ARRA supplement) and S.S. (GM34059). R.G. was supported by a fellowship from the National Science Foundation of Switzerland. L.G. was supported by NIH T32 GM008646. A.D. and K.O. receive salary and other support from the Ludwig Institute for Cancer Research.

Author Contributions R.G., A.M. and T.E. performed ChIP experiments; A.R. performed analysis of all ChIP-chip datasets with advice from S.S.; K.W.Y. performed the photobleaching and α -amanitin experiments; F.B. and P.M. performed the mating inheritance experiment, analysed replication-independence and measured CeCENP-A levels in sperm; R.G. performed GFP–CeCENP-A localization analysis, quantified CeCENP-A levels in nuclei with K.W.Y., and performed qPCR on germline RNA provided by L.G.; A.E., A.M. and J.M. generated GFP–CeCENP-A strains; S.E. and J.D.L. helped initiate ChIP analysis of CeCENP-A; K.O. and A.D. made initial observations that established the project; A.D., R.G., A.R., K.W.Y. and S.S. prepared the figures and wrote the paper with advice from J.D.L. and K.O.; A.D. and S.S. supervised the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.D. (abdesai@ucsd.edu) or S.S. (sstrome@ucsc.edu).

METHODS

Antibodies. ChIP for HCP-3/CeCENP-A was performed with OD79, an affinity-purified rabbit polyclonal antibody raised against amino acids 3–183 of CeCENP-A⁸. A second antibody (SDQ0804) raised against amino acids 40–118 of CeCENP-A confirmed the ChIP-chip pattern observed with antibody OD79 (genome-wide correlation coefficient of 0.92). ChIP for KNL-2 was performed with two polyclonal rabbit antibodies (SDQ08003, SDQ08010) raised against amino acids 207–306. ChIP for RNA Pol II was performed using a monoclonal antibody against the CTD repeat YSPTSPS of RNA Pol II (8WG16; abcam ab817mod). The antibody for H3K36me3 ChIP was described previously¹³. Other antibodies used for immunoblotting and immunofluorescence were: α -tubulin (DM1 α ; Sigma-Aldrich), NPP-112 and SQV-8 (gifts from J. Audhya), BUB-1 (ref. 21) and KNL-2 (ref. 22).

Worm strains. For monitoring GFP-CeCENP-A and GFP-CPAR-1 localization in live adult gonads, sperm and embryos (Fig. 1 and Supplementary Figs 1 and 2), and for the photobleaching experiments of GFP-CeCENP-A in Fig. 1, worm strains expressing transgenes from a single-copy locus that includes endogenous 5' and 3' regions were generated using the MosSCI technique²³, as outlined in Supplementary Fig. 1b. The GFP-CeCENP-A strain OD347 fully rescued the *CeCENP-A/hcp-3* deletion allele *ok1892*, and all experiments involving GFP-CeCENP-A were performed in the *ok1892* deletion background. For imaging, the GFP-CeCENP-A (OD347) and GFP-CPAR-1 (OD588) strains were crossed with a strain expressing mCherry-histone H2b (OD56) to generate strains OD421 and OD416, respectively.

For the mating-based analysis of CeCENP-A inheritance through fertilization (Fig. 1c), wild-type (N2) males were crossed to strain BA17, which harbours a temperature-sensitive mutation that abrogates sperm production. Use of the *fem-1* strain ensured that the embryos analysed by immunofluorescence were cross-progeny. Injection of double-stranded RNA targeting *CeCENP-A* and mating with males was performed as previously described²⁴.

For the transcription inhibition assay (Supplementary Fig. 8), GFP-CeCENP-A-expressing hermaphrodites (OD347) were mated with *his-72p::GFP-H3.3* and *end-3p::mCherry-H1* co-expressing males (RW10007).

For the photobleaching analysis in Supplementary Fig. 4, a strain expressing GFP-CeCENP-A was generated by bombarding plasmid pJM12 into *unc-119(de3)* worms. In pJM12, a GFP-CeCENP-A transgene is expressed under control of the *pie-1* promoter and 3' untranslated region (UTR). The GFP is inserted at amino acid 174 between the amino-terminal tail and histone core (Supplementary Fig. 4a). Coding sequence and introns preceding amino acid 134 were altered to preserve coding information, but make the transgene-encoded mRNA resistant to RNAi. The two introns in this part of the *CeCENP-A* locus were replaced with introns from SPHased GFP (Fire lab 2005 vector kit). Ballistic bombardment of pJM12 generated strain OD136. The transgene insertion in OD136 cannot be homozygosed—there is a low amount of embryonic lethality (12–14%) and Unc progeny. This is probably due to the transgene insertion site, as both of these phenotypes segregated with the GFP fluorescence through multiple outcrosses. A dsRNA to the re-encoded region was used to selectively deplete endogenous CeCENP-A and assess functional rescue by the transgene (Supplementary Fig. 4b). OD136 was used to generate strain OD265, where one copy of the endogenous *CeCENP-A* locus is deleted and GFP-CeCENP-A as well as mCherry-H2b are co-expressed. Photobleaching experiments in Supplementary Fig. 4d, e were performed using OD265. Strain genotypes are listed in Supplementary Table 1.

RNA interference. L4 worms were injected with dsRNA prepared as described previously²¹ and incubated for 48 h at 20 °C, except for the mating experiment using the *fem-1* mutant (see above).

N2 genomic DNA was used as a template to generate PCR products for dsRNA production. The dsRNA for depletion of CeCENP-A in the inheritance experiment of Fig. 1c was described previously⁸.

Oligonucleotides for production of dsRNA against the re-encoded sequence of CeCENP-A used in the rescue experiments with strain OD136 (Supplementary Fig. 4a, b): oOD1887, 5'-AATTAACCCCTCACTAAAGGccgatgacacccaattat-3'; oOD1888, 5'-TAATACGACTCACTATAGGcgtggagtaatcgacaag-3'. Oligonucleotides for dsRNA against GFP: oOD2423, 5'-TAATACGACTCACTATAGGgtcagtggaagggtggaaggtg-3'; oOD2424, 5'-AATTAACCCCTCACTAAAGGcatccatgtgtaatccagcagc-3'.

For replication inhibition (Supplementary Fig. 3), dsRNAs targeting *cdc-6* and *cdt-1* were mixed to obtain equal concentrations. Oligonucleotides used for dsRNA against *cdc-6*: oOD1265, 5'-AATTAACCCCTCACTAAAGGCAAAATTCCTGCTGCTCCAAT-3'; oOD1266, 5'-TAATACGACTCACTATAGGCGGTGCAACCTCAAGTTTCAAT-3'. Oligonucleotides used for dsRNA against *cdt-1*: oOD801, 5'-AATTAACCCCTCACTAAAGGCAAAAACAACGAAGCGTGTG-3';

oOD802, 5'-TAATACGACTCACTATAGGCTCGTTTTTCATTTTATCATTCA-3'.

Immunofluorescence and immunoblotting. Embryos were fixed and processed for immunofluorescence as described previously^{21,25}. Antibodies directly labelled with fluorescent dyes (Cy2, Cy3 or Cy5; Amersham Biosciences) were used at 1 μ g ml⁻¹. Images were recorded on a DeltaVision microscope at 1 \times 1 binning with a \times 100 numerical aperture (NA) 1.3 U-planApo objective (Olympus). Z-stacks (0.2- μ m sections) were deconvolved using softWoRx (Applied Precision), and maximum intensity projections were imported into Adobe Photoshop CS4 for further processing.

Immunoblotting was performed using standard methods. For CeCENP-A immunoblots (Figs 1b and 2e), proteins were transferred for 5 h at 30 V in 25 mM Tris-HCl pH 8.3, 192 mM glycine, 20% methanol. These blotting conditions were optimized to result in quantitative transfer of CeCENP-A onto the membrane.

Live imaging and photobleaching. All live imaging was performed at 20 °C. For images of adult hermaphrodite gonads (Fig. 1a and Supplementary Fig. 2a, b) worms were anesthetized with a mixture of 1 mg ml⁻¹ ethyl 3-aminobenzoate methanesulphonate and 0.1 mg ml⁻¹ of tetramisole hydrochloride in M9 for 15–30 min before transferring them to a 2% agarose pad under a coverslip. Images of gonad regions were acquired with a \times 40 1.3 NA PlanFluor objective by collecting an 80 \times 0.5 μ m Z-series of GFP and mCherry images for every Z-plane. The whole gonad views shown in Fig. 1a and Supplementary Fig. 2a were stitched together from three individual, overlapping images. Embryos (Supplementary Fig. 1c) and -1 oocytes/spermatheca (Supplementary Fig. 2b) were imaged using a \times 100 1.4 NA PlanApoChromat objective. Images were acquired with 1 \times 1 binning on a spinning disk confocal setup mounted on a Nikon TE2000-E inverted microscope equipped with a solid-state laser combiner (ALC) (491 nm and 561 nm lines), a Yokogawa CSU10 head and a CCD camera (iXon DV887; Andor Technology). Acquisition parameters, shutters and focus were controlled by iQ 1.10.0 software (Andor Technology). Images were processed with Fiji 1.0 and Adobe Photoshop CS4.

For the transcription inhibition assay (Supplementary Fig. 8), cross-progeny embryos from GFP-CeCENP-A expressing hermaphrodites mated with *his-72p::GFP-H3.3* and *end-3p::mCherry-H1* co-expressing males were dissected in L-15 blastomere culture medium²⁶ containing 200 μ g ml⁻¹ α -amanitin. GFP and mCherry Z-stacks of permeable and impermeable embryos in the same field of view were acquired at 1–4 min intervals with a \times 60 1.4 NA PlanApoChromat objective until embryos contained more than 50 cells.

The photobleaching experiments in Fig. 1d, e were performed with the FRAPPA unit (Andor Technology) using a \times 60 1.4 NA PlanApoChromat objective. Thirteen Z-sections were acquired with a spacing of 1 μ m before and after bleaching in the first embryonic division and at 1-min intervals thereafter until anaphase of the second division. Maximum intensity projections were generated for each Z-stack. For each image sequence, an identical sized rectangle (R1) was drawn around each anaphase chromatid set before and after photobleaching in the first division and around anaphase chromatid sets in the second division. A larger rectangle (Rb) was drawn around each rectangle R1 and the area between the two rectangles served as a measure of background intensity. The average intensity (Avg. Int.) of the GFP signal in each R1 was measured using the formula: Avg. Int._{R1} = [(Avg. Int._{Rb} \times Area_{Rb} - Avg. Int._{R1} \times Area_{R1}) / (Area_{Rb} - Area_{R1})], and the ratio of average intensities on anaphase chromatid sets before/after bleaching and in the subsequent anaphase were calculated and averaged.

For the photobleaching experiments in Supplementary Fig. 4d, e, the microscope setup differed from the one used for the experiments in Fig. 1 as follows: the microscope was equipped with a krypton-argon 2.5 W water-cooled laser (Spectra-Physics), acquisition parameters, shutters and focus were controlled by MetaMorph software (MDS Analytical Technologies), and the 488 nm laser line for photobleaching was steered into a custom-modified epifluorescence port. GFP-CeCENP-A intensity ratios were calculated as described above, except that anaphase chromatid sets before/after bleaching in the first division were compared with metaphase plates in the second division.

Expression analysis on dissected germlines by quantitative PCR. Worms containing one or two embryos were dissected with 30-gauge needles in Egg buffer (25 mM HEPES-KOH pH 7.6, 118 mM NaCl, 48 mM KCl, 2 mM CaCl₂, 2 mM MgCl₂) containing 1 mM levamisole and 0.5% Tween-20. Total RNA from 50–100 gonads was isolated using TRIzol (Invitrogen), and cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen). Quantitative real-time PCR was performed with iQ SYBR Green Supermix (Bio-Rad) in the iQ5 cyclor (Bio-Rad) using standard protocols. The average amplification efficiency (*E*) of primer pairs (Supplementary Table 2) was calculated from two standard curves (tenfold dilution series of cDNA prepared from mixed-stage N2 worms). The relative transcript abundance (RTA) of target genes in *met-1* mutant (*met-1*)

versus wild-type (WT) germlines was calculated after normalization to the reference gene *act-2* (actin homologue), using the formula: $RTA = \frac{E_{\text{target}}^{(C_{\text{WT}} - C_{\text{met-1}})}}{E_{\text{reference}}^{(C_{\text{WT}} - C_{\text{met-1}})}}$,

where C_t denotes the threshold cycle. The assay was performed in duplicate on four biological replicates each for wild-type and *met-1* mutant germlines.

Quantification of CeCENP-A molecules in purified nuclei and sperm. 6xHis-CeCENP-A was expressed in bacteria, purified by nickel affinity chromatography under denaturing conditions followed by electroelution using the Bio-Rad Electro-Eluter. The concentration of the purified protein was measured relative to BSA on a gel. Purified sperm were a gift of S. Ward. Sperm concentration was measured by microscopy following 4',6'-diamidino-2-phenylindole (DAPI) staining.

For isolation of nuclei, early embryos (<100 cells) were harvested from synchronized adult worms and treated with chitinase as described below for embryo extract preparation. Packed embryos (1 ml) were washed with 2×50 ml chilled Egg buffer (25 mM HEPES-KOH pH 7.6, 118 mM NaCl, 48 mM KCl, 2 mM CaCl_2 , 2 mM MgCl_2), hypotonically swollen for 15 min in 10 ml of $0.5 \times$ Nuclei buffer (5 mM Tris-HCl pH 8, 40 mM KCl, 1 mM K-EDTA, 0.375 mM spermidine, 0.15 mM spermine), then washed into 10 ml of $1 \times$ Nuclei buffer (10 mM Tris-HCl pH 8, 80 mM KCl, 2 mM K-EDTA, 0.75 mM spermidine, 0.3 mM spermine) supplemented with 0.1% digitonin (Sigma-Aldrich) and protease inhibitors, and immediately dounced with about 50 strokes in a 15-ml Wheaton Dounce homogenizer using pestle B. Large debris was pelleted at 100g for 3 min and re-dounced once as above. Supernatants containing the nuclei were combined and spun at 2,000g for 15 min. The nuclei pellet was suspended in $1 \times$ Nuclei buffer supplemented with 0.1% digitonin and protease inhibitors and layered onto a 30% (w/v) in Nuclei buffer + 0.1% digitonin) sucrose cushion. Nuclei were recovered in the pellet after spinning at 2,000g for 15 min.

Embryo isolation, fixation and extract preparation. N2 adult worms were grown from synchronized L1 larvae in S-basal medium. Batches of 500 ml in 2.8-l Fernbach flasks shaking at 230 r.p.m. were incubated at 17 °C (early embryos) or 19 °C (late embryos) for approximately 65 h. The exact time of harvest was determined by checking embryo production under a microscope (for early embryos, this was five embryos per worm or less). Gravid adults were separated from debris by sucrose floating, and embryos were recovered by dissolving adults with a bleach/NaOH solution. 10 μ l and 2 μ l of packed embryos were set aside for expression profiling and staging by fluorescence microscopy after DAPI staining, respectively. The remainder was suspended in 2 volumes of Egg buffer (25 mM HEPES-KOH pH 7.6, 118 mM NaCl, 48 mM KCl, 2 mM CaCl_2 , 2 mM MgCl_2) and incubated with 0.15 units ml^{-1} chitinase (Sigma-Aldrich) until visible disintegration of the eggshell. Embryos were washed with 2×50 ml chilled phosphate-buffered saline (PBS) and suspended in 40 ml PBS. Fixation was performed for 10 min on ice after adding 4 ml of cross-linking solution (11% formaldehyde, 50 mM HEPES-KOH pH 8, 0.1 M NaCl, 1 mM Na-EDTA, 0.5 mM Na-EGTA), and excess formaldehyde was quenched with 120 mM glycine. Fixed embryos were washed with 3×50 ml PBS and suspended in five pellet volumes of ChIP buffer (50 mM HEPES-KOH pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.5% NP-40, protease inhibitors). Sonication was performed using a Branson sonifier microtip in cycles of 10 s duration (0.9 s on, 0.1 s off) with the power setting at 25% (2 cycles), 30% (2 cycles), 35% (10 cycles), 40% (2 cycles), and 45% (2 cycles), and a pause of 1 min between cycles. Crude extracts were spun for 20 min at 10,000g, the supernatant was removed and glycerol added to 10%. Protein concentration was determined by the Bradford method and aliquots of 3 mg protein were flash-frozen in liquid nitrogen.

Chromatin immunoprecipitation (ChIP). Extract corresponding to 3 mg protein was diluted to 900 μ l with ChIP buffer. After addition of sarcosyl to 1%, Na-deoxycholate to 0.1%, and PMSF to 1 mM, the extract was spun for 10 min at maximum speed in a tabletop centrifuge and the supernatant was used for ChIP. 50 μ l was removed for preparation of input DNA. To the rest, 50 μ l of Dynabead Sheep anti-Rabbit or anti-Mouse IgG suspension (Dyna Beads), pre-bound to 5 μ g of target antibody, were added, and the mixture was incubated at 4 °C for 4 h or overnight (<16 h). Beads were recovered with a Dynal Magnetic Particle Concentrator (Invitrogen) and washed 2×5 min with buffer FA (50 mM HEPES-KOH pH 7.6, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate), 10 min with FA-1000 (50 mM HEPES-KOH pH 7.6, 1 M NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate), 10 min with FA-500 (50 mM HEPES-KOH pH 7.6, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) in a new tube, 10 min with TEL buffer (10 mM Tris-HCl pH 8.0, 0.25 M LiCl, 1 mM EDTA, 1% NP-40, 1% Na-deoxycholate), and once briefly with TE (10 mM Tris pH 8, 1 mM EDTA). Antibody-bound chromatin was recovered in 50 μ l Elution buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 250 mM NaCl, 1% SDS) by shaking at 67 °C for 15 min. Input and ChIP chromatin samples were subsequently processed in parallel. Cross-links were reversed over night at 65 °C in Elution buffer, proteins were digested with 0.45 mg

ml^{-1} proteinase K for 2 h at 37 °C, and nucleic acids were recovered by phenol/chloroform extraction and precipitation with ethanol. RNA was digested with 0.3 mg ml^{-1} RNase A in TE at 37 °C for 2 h, and DNA purified with a column (PCR Purification Kit, Qiagen).

The ChIP procedure for the *met-1* mutant analysis (Fig. 4 and Supplementary Figs 11 and 12) differs slightly from the above and was described in detail previously¹³.

Ligation-mediated PCR. DNA ends were blunted with 5 units ml^{-1} T4 DNA polymerase (New England Biolabs) at 12 °C for 20 min and the DNA recovered by phenol/chloroform extraction and ethanol precipitation. Annealed oligomer adaptors (oligo1: 5'-GCGGTGACCCGGGAGATCTGAATTC-3'; oligo2: 5'-GAATTCAGATC-3') were ligated to blunt DNA ends at 2 μ M with 4,000 units ml^{-1} T4 DNA ligase (New England Biolabs) overnight at 16 °C, and DNA was precipitated with ethanol. DNA fragments were amplified for 22 cycles (55 °C, 2 min; 72 °C, 5 min; 95 °C, 2 min; 95 °C, 1 min; 60 °C, 1 min; 72 °C, 2 min; start cycle again at step 4) using 1 μ M oligo 1 and a *Taq* DNA polymerase (100 units ml^{-1})/*Pfu* DNA polymerase (0.5 units ml^{-1}) mix. Amplified DNA was purified with a column (PCR Purification Kit, Qiagen).

Microarray hybridizations, data analysis and display. Amplified ChIP DNA was labelled and hybridized by the Roche Nimblegen Service Laboratory. 2.1M probe tiling arrays, with 50-bp probes, designed against WormBase version WS170 (ce4) were used for all experiments. ChIP samples were labelled with Cy5 and their input reference with Cy3. One ChIP was dye-swapped, which resulted in the same pattern (not shown). For each probe, the intensity from the sample channel was divided by the reference channel and transformed to log. The enrichment scores for each replicate were calculated by standardizing the log ratios to mean zero and standard deviation one (*z*-score). Genome-wide scatter plots and Pearson correlations between all ChIP targets and replicates were obtained using all probe *z*-scores after median smoothing over 1-kb windows.

The average *z*-score of two replicates was used for all analyses, except in Supplementary Figs 9, 10 and 13c, where individual data sets from extracts with distinct age distributions of embryos are compared. Accession numbers for data sets used in this study are listed in Supplementary Table 3. Scatter plots and boxplots for genes were generated by averaging *z*-scores of probes located completely within the transcript start site (TSS) and end site (TES). TSS and TES coordinates were obtained from WormBase (WS170).

Gene body profile plots (Supplementary Fig. 7b) were generated by aligning genes of length greater than 2 kb at their TSS and TES. The genomic regions 1.5 kb upstream to 1 kb downstream from TSS and 1 kb upstream to 1 kb downstream from TES were divided into 50-bp bins, and probes were assigned to the nearest bin. Gene group profiles were generated by averaging probe *z*-scores within each bin across genes in the group.

Definition of CeCENP-A-enriched domains. CeCENP-A signal was averaged over 2-kb windows, every 50 bp. A random distribution for the window averages was obtained by randomly sampling and assigning CeCENP-A values for each chromosome. The resulting random CeCENP-A tracks were also averaged over 2-kb windows. A cutoff was selected so that the number of random window averages above the cutoff was less than 3% of the number of non-randomized windows above the cutoff, effectively providing a 3% false positive rate with respect to the random window averages. Overlapping windows above the cutoff were combined into domains. Domains with gaps smaller than 2 kb were merged, and domains smaller than 2.5 kb were excluded.

Definition of gene classes based on expression profiling data sets. Gene classes were defined on the basis of expression data, as described previously¹³. In brief, 'Ubiquitous' or housekeeping have transcripts present in muscle, gut, neuron and adult germline SAGE (serial analysis of gene expression) data sets^{27,28}; 'Germline-expressed' genes have transcripts present in the dissected adult hermaphrodite germ line SAGE data set²⁷; 'Serpentine' receptor genes are expressed in mature neurons and silent in embryos²⁹; 'Spermatogenesis' genes are classified as expressed during sperm production on the basis of comparative microarray analysis³⁰; 'Germline-only' genes are expressed exclusively in the maternal germ line, as their transcripts are enriched in the germline³⁰, maternally loaded into embryos¹¹, and absent from muscle, gut and neuron SAGE data sets^{27,28}; transcripts of 'Embryo-expressed' genes are not maternally provided and increase in level during embryogenesis¹¹.

Criteria for identifying genes with maximal changes in RNA Pol II levels. Genes with maximal changes in RNA Pol II levels between the two averaged early embryo (EE) and late embryo (LE) extracts (Supplementary Fig. 10a, b), were identified by applying a moderated *t*-test³¹ and requiring a false discovery rate smaller than 5% (ref. 32). In addition, RNA Pol II levels for those genes were required to show at least a twofold change in RNA Pol II ChIP-chip hybridization signal between the averaged EE and LE extracts.

Transcriptional profiling of embryos. RNA was isolated from 10 μ l of packed embryos using TRIzol (Invitrogen) and the RNeasy kit (Qiagen). RNA (20 μ g) was

hybridized to a single-colour 4-plex Nimblegen expression array with 72,000 probes (three 60-mer oligo probes per gene). Quantile normalization³³ and the robust multichip average (RMA) algorithm³⁴ were used to normalize and summarize the multiple probe values per gene to obtain one expression value per gene and sample. The expression values per gene were averaged across samples as indicated in the figure legends.

21. Oegema, K., Desai, A., Rybina, S., Kirkham, M. & Hyman, A. A. Functional analysis of kinetochore assembly in *Caenorhabditis elegans*. *J. Cell Biol.* **153**, 1209–1226 (2001).
22. Maddox, P. S., Hyndman, F., Monen, J., Oegema, K. & Desai, A. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J. Cell Biol.* **176**, 757–763 (2007).
23. Frøkjaer-Jensen, C. *et al.* Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nature Genet.* **40**, 1375–1383 (2008).
24. Dammermann, A. *et al.* Centriole assembly requires both centriolar and pericentriolar material proteins. *Dev. Cell* **7**, 815–829 (2004).
25. Cheeseman, I. M. *et al.* A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **18**, 2255–2268 (2004).
26. Edgar, L. G. Blastomere culture and analysis. *Methods Cell Biol.* **48**, 303–321 (1995).
27. Wang, X. *et al.* Identification of genes expressed in the hermaphrodite germ line of *C. elegans* using SAGE. *BMC Genomics* **10**, 213 (2009).
28. Meissner, B. *et al.* An integrated strategy to study muscle development and myofilament structure in *Caenorhabditis elegans*. *PLoS Genet.* **5**, e1000537 (2009).
29. Kolasinska-Zwierz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature Genet.* **41**, 376–381 (2009).
30. Reinke, V., Gil, I. S., Ward, S. & Kazmer, K. Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. *Development* **131**, 311–323 (2004).
31. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3 (2004).
32. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).
33. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
34. Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).

Hsp72 preserves muscle function and slows progression of severe muscular dystrophy

Stefan M. Gehrig¹, Chris van der Poel^{1†}, Timothy A. Sayer¹, Jonathan D. Schertzer^{1†}, Darren C. Henstridge², Jarrod E. Church^{1†}, Severine Lamon³, Aaron P. Russell³, Kay E. Davies⁴, Mark A. Febbraio² & Gordon S. Lynch¹

Duchenne muscular dystrophy (DMD) is a severe and progressive muscle wasting disorder caused by mutations in the dystrophin gene that result in the absence of the membrane-stabilizing protein dystrophin^{1–3}. Dystrophin-deficient muscle fibres are fragile and susceptible to an influx of Ca^{2+} , which activates inflammatory and muscle degenerative pathways^{4–6}. At present there is no cure for DMD, and existing therapies are ineffective. Here we show that increasing the expression of intramuscular heat shock protein 72 (Hsp72) preserves muscle strength and ameliorates the dystrophic pathology in two mouse models of muscular dystrophy. Treatment with BGP-15 (a pharmacological inducer of Hsp72 currently in clinical trials for diabetes) improved muscle architecture, strength and contractile function in severely affected diaphragm muscles in *mdx* dystrophic mice. In *dko* mice, a phenocopy of DMD that results in severe spinal curvature (kyphosis), muscle weakness and premature death^{7,8}, BGP-15 decreased kyphosis, improved the dystrophic pathophysiology in limb and diaphragm muscles and extended lifespan. We found that the sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase (SERCA, the main protein responsible for the removal of intracellular Ca^{2+}) is dysfunctional in severely affected muscles of *mdx* and *dko* mice, and that Hsp72 interacts with SERCA to preserve its function under conditions of stress, ultimately contributing to the decreased muscle degeneration seen with Hsp72 upregulation. Treatment with BGP-15 similarly increased SERCA activity in dystrophic skeletal muscles. Our results provide evidence that increasing the expression of Hsp72 in muscle (through the administration of BGP-15) has significant therapeutic potential for DMD and related conditions, either as a self-contained therapy or as an adjuvant with other potential treatments, including gene, cell and pharmacological therapies.

DMD is the most severe form of muscular dystrophy; it affects about 1 in 3,500 live male births¹. Intracellular Ca^{2+} regulation is compromised in dystrophic muscle fibres, which triggers chronic inflammation, repeated cycles of degeneration with progressively ineffective regeneration, and infiltration of fibrotic and other non-contractile material⁴. Mechanisms for the influx of Ca^{2+} into dystrophic muscle fibres include membrane tears^{4–6}, stretch-activated channels⁹, Ca^{2+} leak channels¹⁰ and leaky Ca^{2+} release channels¹¹, and it has been speculated that the function of SERCA, the main protein responsible for Ca^{2+} reuptake into the sarcoplasmic reticulum (SR), is compromised^{12,13}. Increasing SERCA pump expression within dystrophic muscles in transgenic mice or through viral-mediated delivery improves Ca^{2+} handling and suppresses the pathological cascade of events^{14,15}. The role of inflammation in the dystrophic pathology is well known, particularly that of the pro-inflammatory cytokine tumour necrosis factor- α (TNF- α)¹⁶. TNF- α activates the nuclear factor- κB (NF- κB) and c-Jun N-terminal kinase (JNK) signalling pathways^{17–19}. Hsp72 is a molecular chaperone protein that inhibits inflammatory mediators including p-JNK, TNF- α and the

NF- κB pathway^{20–22}, and binds and preserves SERCA function under conditions of cellular stress²³. Although some studies have shown Hsp72 to be elevated in patients with DMD²⁴, there is little consensus, because expression data for young patients are variable and sourcing age-matched controls is problematic (Supplementary Fig. 1a). Nevertheless, the endogenous heat shock protein response in DMD is insufficient to be protective. Here we tested the hypothesis that increasing the levels of Hsp72 protects dystrophic muscles from functional deterioration. We bred dystrophin-null *mdx* dystrophic mice with mice showing a muscle-specific transgenic (TG) overexpression of Hsp72, producing *mdx*^{TG(+)} mice and *mdx* littermate controls (Fig. 1a; see Supplementary Fig. 1b for quantification). At about 25 weeks of age, serum levels of creatine kinase (CK), a classic indicator of muscle breakdown, were decreased in *mdx*^{TG(+)} mice compared with littermate control mice lacking the transgene (*mdx*^{TG(–)}; Fig. 1b). Because most patients with DMD show severe weakness⁴ and/or muscle fatigue²⁵, we assessed whole-body strength and endurance in dystrophic mice by using a hang test to measure latency-to-fall, which was significantly improved in *mdx*^{TG(+)} mice (24 ± 3 s versus 64 ± 14 s; $P = 0.002$, $n \geq 20$ mice). Respiratory failure is the cause of death in up to 90% of patients with DMD²⁶, and because diaphragm function is an accurate predictor of respiratory insufficiency we investigated the effect of Hsp72 overexpression on the pathophysiology of the diaphragm in dystrophin-deficient mice. The progressive degeneration of the diaphragm in *mdx* mice closely mimics that in DMD²⁷. Gross histological analyses revealed that the diaphragm pathology in *mdx*^{TG(+)} mice was ameliorated compared with age-matched littermate control *mdx*^{TG(–)} mice (Supplementary Fig. 1c), an observation supported by the minimal Feret's diameter variance coefficient, which provides a sensitive measure of fibre heterogeneity and the dystrophic pathology²⁸. We found a lower Feret's diameter variance coefficient in *mdx*^{TG(+)} mice than in *mdx*^{TG(–)} mice, indicative of an improved phenotype (405 ± 5 versus 358 ± 5 ; $P < 0.001$, $n = 5$). Damaged myofibres can be revealed by the infiltration of Evans blue dye (EBD) entering the myoplasm through tears in the sarcolemma^{4–6}. EBD infiltration was decreased in cross-sections of the diaphragm from *mdx*^{TG(+)} mice in comparison with that from *mdx*^{TG(–)} mice (Fig. 1c and Supplementary Fig. 1d), indicative of decreased necrosis and hastened overall repair of damaged fibres rather than improved structural integrity. To support this contention, we performed well-described *in situ* and *in vitro* contraction-induced injury protocols on tibialis anterior (TA) and diaphragm muscles, respectively (see Methods). No differences in contraction-mediated damage were evident between muscles of *mdx*^{TG(+)} and *mdx*^{TG(–)} mice (Supplementary Fig. 1e), indicating that structural integrity was unaltered. Expression of the dystrophin homologue, utrophin, a protein known to compensate for the loss of dystrophin, was also unchanged (Supplementary Fig. 1f). We assessed collagen infiltration in sections of

¹Basic and Clinical Myology Laboratory, Department of Physiology, University of Melbourne, Victoria, 3010, Australia. ²Cellular and Molecular Metabolism Laboratory, Baker-IDI Heart and Diabetes Institute, PO Box 6492, St Kilda Road Central, Victoria, 8008, Australia. ³Centre for Physical Activity and Nutrition Research, School of Exercise and Nutrition Sciences, Deakin University, Burwood, Victoria, 3125, Australia. ⁴MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK. [†]Present addresses: Department of Human Biosciences, Faculty of Health Sciences, La Trobe University, Bundoora, 3086, Victoria, Australia (C.v.d.P. and J.E.C.); Department of Biochemistry and Biomedical Sciences and Department of Pediatrics, McMaster University, Hamilton, Ontario, L8S 4L8, Canada (J.D.S.).

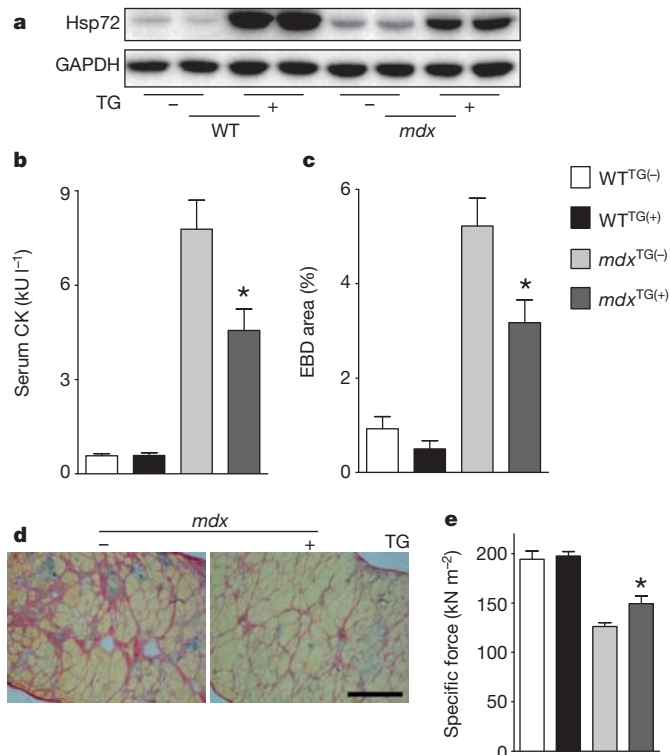


Figure 1 | Transgenic Hsp72 overexpression increases muscle strength, decreases muscle breakdown and improves diaphragm muscle histological parameters in *mdx* mice. **a**, Representative western blot detection of Hsp72 in diaphragm muscle homogenates from non-dystrophic (WT) and dystrophic (*mdx*) Hsp72 transgene-negative (TG⁻) and transgene-positive (TG⁺) mice. GAPDH, glyceraldehyde-3-phosphate dehydrogenase. **b**, Whole-body muscle breakdown, measured by serum creatine kinase (CK) levels. Asterisk, $P < 0.05$, $n \geq 6$. **c**, Quantified mean data for EBD-positive area in diaphragm muscle sections. Asterisk, $P < 0.05$, $n \geq 4$. **d**, Representative images of collagen infiltration in diaphragm sections (revealed with Van Gieson's stain) from *mdx*^{TG(-)} and *mdx*^{TG(+)} mice. **e**, Specific (normalized) force of diaphragm muscle strips measured *in vitro*. Asterisk, $P < 0.05$, $n \geq 5$. All data are from 25-week-old mice. Scale bar, 200 μ m. Data are shown as means \pm s.e.m.

diaphragm and found that *mdx*^{TG(+)} mice had less collagen infiltration than *mdx*^{TG(-)} mice at both 25 weeks of age (Fig. 1d and Supplementary Fig. 1g) and 80 weeks of age (data not shown). Normalized force production was significantly higher in diaphragm muscle strips from *mdx*^{TG(+)} mice than in those from *mdx*^{TG(-)} mice (Fig. 1e).

We have shown that Hsp72 can block inflammation in mice *in vivo*²², and because inflammation promotes muscle degeneration in mouse models of DMD¹⁷ we examined the effect of Hsp72 on expression of p-JNK, p-IKK (a key mediator of NF- κ B activation) and TNF- α . Western blot and polymerase chain reaction (PCR) analyses revealed no difference in these inflammatory markers in diaphragm muscles from *mdx*^{TG(-)} and *mdx*^{TG(+)} mice (Supplementary Fig. 2, and data not shown). There was significantly decreased messenger RNA expression of macrophage markers CD68 and F4/80 and TNF- α in TA muscles of *mdx*^{TG(+)} mice (Supplementary Fig. 3), but these did not translate to functional improvements (data not shown). We next examined SERCA activity in diaphragm muscles from wild-type (WT) and *mdx* mice and showed a progressive age-related decline in maximal SERCA activity in *mdx* mice (Fig. 2a and Supplementary Fig. 4a), despite an age-related increase in SERCA protein expression (Supplementary Fig. 4b). This functional decline is attributed, in part, to post-translational modifications of the SERCA protein, especially nitrosylation, which decreases maximal SERCA activity as a result of changes in Ca²⁺-binding and ATP-binding domains²³. Alterations in the Ca²⁺-binding domain changes SERCA

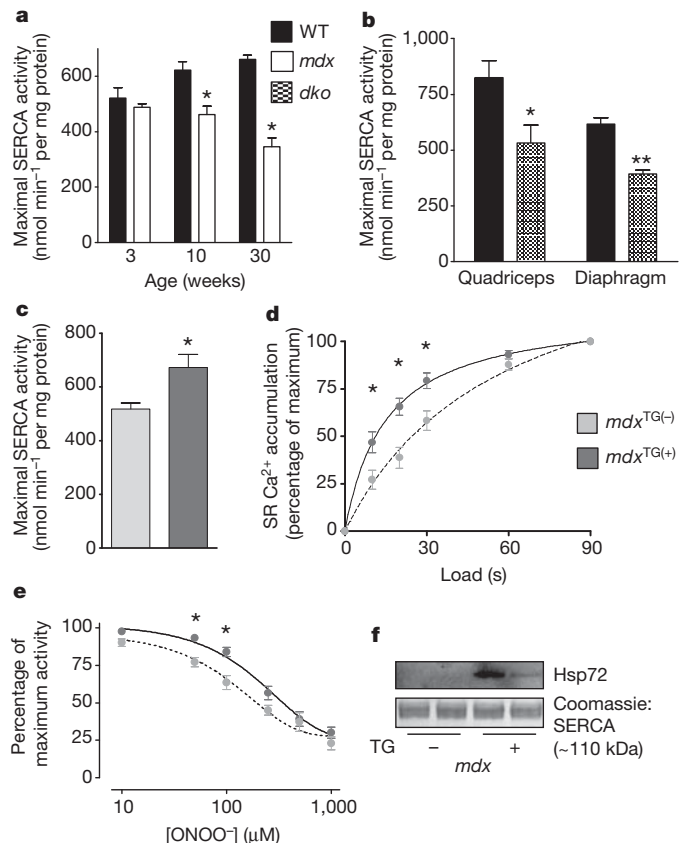


Figure 2 | Maximal SERCA activity is decreased in mouse models of dystrophy; Hsp72 binding improves SERCA function. **a**, Maximal SERCA activity in diaphragm muscle homogenates from WT and *mdx* mice at 3, 10 and 30 weeks of age. Asterisk, $P < 0.001$, $n \geq 4$. **b**, Maximal SERCA activity of muscle homogenates from quadriceps and diaphragm from 10-week-old WT and *dko* mice. Asterisk, $P = 0.036$. Two asterisks, $P < 0.001$, $n \geq 4$. **c**, Maximal SERCA activity in muscle homogenates from *mdx*^{TG(-)} (pale grey bars) and *mdx*^{TG(+)} (dark grey bars) mice. Asterisk, $P < 0.008$, $n \geq 9$. **d**, Ca²⁺ accumulation curves (an indirect measure of SERCA activity) for SR in single fibres from extensor digitorum longus muscles of *mdx*^{TG(-)} and *mdx*^{TG(+)} mice. Asterisk, $P < 0.01$, $n \geq 15$. **e**, Maximal SERCA activity after incubation for 5 min at various concentrations of peroxynitrite (ONOO⁻), measured in enriched SR vesicles isolated from muscles of *mdx*^{TG(-)} and *mdx*^{TG(+)}. SERCA activity was normalized to maximal activity in the absence of ONOO⁻. Asterisk, $P < 0.05$, $n \geq 3$. **f**, Representative western blots of enriched SR vesicles from *mdx*^{TG(-)} and *mdx*^{TG(+)} ($n = 4$) mice, showing Hsp72 protein expression and Coomassie blue stain (showing SERCA isoforms). Data are shown as means \pm s.e.m.

enzyme kinetics, decreasing Ca²⁺ sensitivity and increasing [Ca²⁺]₅₀ (the [Ca²⁺] required to achieve half-maximal enzyme activity), an effect we observed in diaphragm muscles from 30-week-old *mdx* mice (Supplementary Fig. 4c). Similar deficits in SERCA activity were evident in both limb and diaphragm muscles of severely affected *dko* mice (Fig. 2b), indicating that abnormal SERCA function may contribute to the disruptions in Ca²⁺ regulation characteristic of dystrophic muscles. Indeed, recent evidence suggests that closer regulation of Ca²⁺ homeostasis through enhanced SERCA expression or activity significantly suppresses the degeneration of dystrophic muscle^{14,15}. Because Hsp72 binds SERCA and prevents functional inactivation under conditions of cellular stress²³, we next tested the hypothesis that overexpression of Hsp72 in dystrophic muscles would improve SERCA function. We examined maximal SERCA activity in homogenates of quadriceps muscles, and found an increase in activity in *mdx*^{TG(+)} compared with *mdx*^{TG(-)} mice (Fig. 2c). To support this finding we also examined Ca²⁺ accumulation in SR (an indirect measure of SERCA activity) in single muscle fibres dissected from fast-twitch

extensor digitorum longus, and predominantly slow-twitch soleus muscles, and found a similar increase in maximal SERCA activity in individual muscle fibres in $mdx^{TG(+)}$ compared with $mdx^{TG(-)}$ mice (as for whole-muscle homogenates), with no changes in mRNA for SERCA or in protein expression (Fig. 2d, Supplementary Fig. 5 and Supplementary Table 1). We then examined whether Hsp72 overexpression could protect SERCA function under conditions of stress, like that induced by reactive oxygen–nitrogen species such as peroxynitrite ($ONOO^-$). We developed a SERCA activity assay in which enriched SR vesicles isolated from $mdx^{TG(-)}$ and $mdx^{TG(+)}$ mice were incubated with increasing concentrations of $ONOO^-$; the subsequent suppression in activity was normalized to that in the absence of $ONOO^-$. We found that SERCA activity was greater in the presence of various $ONOO^-$ concentrations in enriched SR vesicles from $mdx^{TG(+)}$ compared with $mdx^{TG(-)}$ mice (Fig. 2e), with western blots revealing Hsp72 protein levels were highly elevated in enriched SR vesicles from $mdx^{TG(+)}$ mice (Fig. 2f; see Supplementary Fig. 6 for full blots), indicating that Hsp72 was bound within the SR to mediate this protective effect. These data indicate that Hsp72 overexpression in dystrophic muscles can protect SERCA from inactivating modifications and is a likely mechanism of protection in $mdx^{TG(+)}$ mice.

Given the significant phenotypic improvements in $mdx^{TG(+)}$ mice, especially in the diaphragm, we examined whether similar effects could be achieved through the pharmacological or heat-therapy induction of Hsp72. BGP-15 is a pharmacological inducer of Hsp72 that can protect against obesity-induced insulin resistance²² and is in Food and Drug Administration (FDA)-approved Phase II clinical trials for diabetes²⁹. Dystrophic mdx mice were treated from 4 to 9 weeks of age (5 weeks) or from 4 to 16 weeks of age (12 weeks) with BGP-15 (15 mg kg⁻¹ per day, oral gavage). WT mice were used for comparisons with mdx mice. Hsp72 protein expression was elevated in diaphragm muscle homogenates from BGP-15-treated mdx mice (and heat-therapy-treated mice) compared with untreated mdx mice (Fig. 3a; see Supplementary Fig. 7a for quantification). CK levels were decreased in BGP-15-treated mdx mice compared with untreated mdx mice (Fig. 3b). EBD infiltration was also reduced with long-term BGP-15 treatment (Fig. 3c). Strength and endurance was evaluated with the inverted hang test; WT mice were stronger than mdx mice, and BGP-15-treated mdx mice showed an increased latency-to-fall compared with control (Supplementary Fig. 7). An increase in fibrosis, as seen in DMD⁴, was observed in the mdx diaphragm (compared with WT), and BGP-15 treatment significantly decreased fibrosis (Fig. 3d; see Supplementary Fig. 7 for quantification). Treatment with BGP-15 attenuated the functional deterioration of the diaphragm muscle significantly (Fig. 3e). Maximal SERCA activity in diaphragm homogenates was increased in mdx mice after long-term treatment with BGP-15, indicating a mechanism consistent with that of transgenic Hsp72 overexpression (Fig. 3f). Because elevated core temperature is a potent inducer of heat shock proteins^{22,30}, we also extensively tested this method of heat shock protein induction (see Methods). Similar beneficial effects to those observed with transgenic Hsp72 overexpression and BGP-15 treatment were seen in mdx mice exposed to repeated heat therapy (see Supplementary Fig. 8).

We then investigated whether treatment with BGP-15 was protective in severely affected dystrophic dco mice, the most phenotypically accurate murine model of DMD^{7,8} (see Methods). The dco mice were treated with BGP-15 from 3–4 weeks until 10 weeks of age. For the survival study, dco mice were treated from 3–4 weeks until death or humane killing (as described in Methods). Photographs of control and BGP-15-treated dco mice were taken at 10 weeks, immediately before killing, and after evisceration and staining of the skeleton with alizarin red (Fig. 4a). Data from WT mice (as in Fig. 3) were also used for comparisons with dco mice. Because boys with DMD have significant paraspinal muscle weakness and in many cases severe kyphosis (spinal curvature)^{4,8}, this was quantified in conscious mice (from the 10-week endpoint cohort) by a blinded investigator using a 1–5 index of

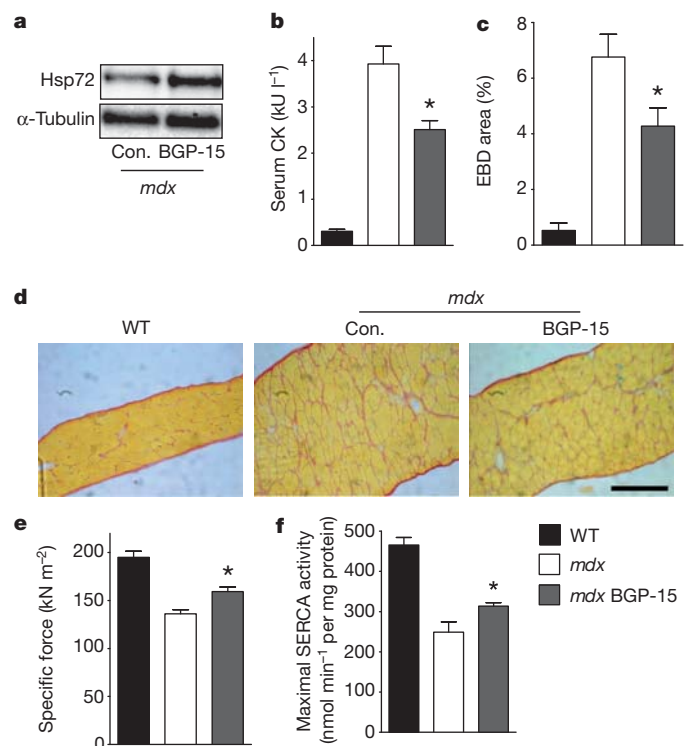


Figure 3 | Pharmacological induction of Hsp72 ameliorates muscular dystrophy in mdx mice. **a**, Representative western blot detection of Hsp72 in diaphragm muscle homogenates of control (Con.) and BGP-15-treated mdx mice. **b**, Whole-body muscle breakdown, measured by serum CK levels in WT (C57BL/10), mdx (control) and BGP-15-treated mdx mice. Asterisk, $P < 0.01$, $n \geq 8$. **c**, Quantified mean data for EBD infiltration in mdx mice treated for 12 weeks. Asterisk, $P < 0.05$, $n \geq 8$. **d**, Representative images of collagen infiltration in diaphragm muscle sections. **e**, Specific force production in diaphragm muscle strips measured *in vitro*. Asterisk, $P < 0.01$, $n \geq 9$. **f**, Maximal SERCA activity in diaphragm muscle homogenates from mdx mice treated for 12 weeks. Asterisk, $P < 0.05$, $n \geq 6$. Scale bar, 200 μ m. WT data are used as a reference control. All treated mice received BGP-15 for 5 weeks unless stated otherwise. Data are shown as means \pm s.e.m.

kyphosis, 1 indicating no spinal deformity on palpation and 5 being the most severe. Treatment with BGP-15 decreased kyphosis markedly compared with vehicle-treated controls (Fig. 4b). Serum CK levels were significantly lower in dco mice after treatment with BGP-15 (Fig. 4c), indicating decreased whole-body muscle breakdown. Collagen infiltration was decreased in the diaphragm of dco mice after treatment with BGP-15 (Fig. 4d) and the force-producing capacity of diaphragm muscle strips and intact TA muscles (measured *in situ*) was increased significantly in BGP-15-treated dco mice, with maximum force restored to WT levels in the TA muscle (Fig. 4e, f and Supplementary Table 2). No differences in body mass or in calcification or central nucleation within diaphragm muscles were evident in dco mice after treatment with BGP-15 (data not shown). However, the most important outcome was that lifelong treatment of dco mice with BGP-15 significantly extended survival (Fig. 4g, h; $P < 0.05$; 27% increase in median lifespan). This finding has clinical relevance for DMD.

Our findings reveal that transgenic Hsp72 overexpression improves several pathological indices in mdx dystrophic mice, at least in part by preserving or improving SERCA function. Furthermore, treatment of dystrophic mice with BGP-15, a known pharmacological co-inducer of Hsp72, ameliorated the dystrophic pathology and extended the lifespan in dco mice. Taken together, these results indicate that induction of Hsp72 in muscular dystrophy is an important and novel therapeutic approach that can improve the dystrophic pathology and attenuate the disease progression. Although an ultimate cure for DMD is likely to be derived from gene or cell therapies, considerable

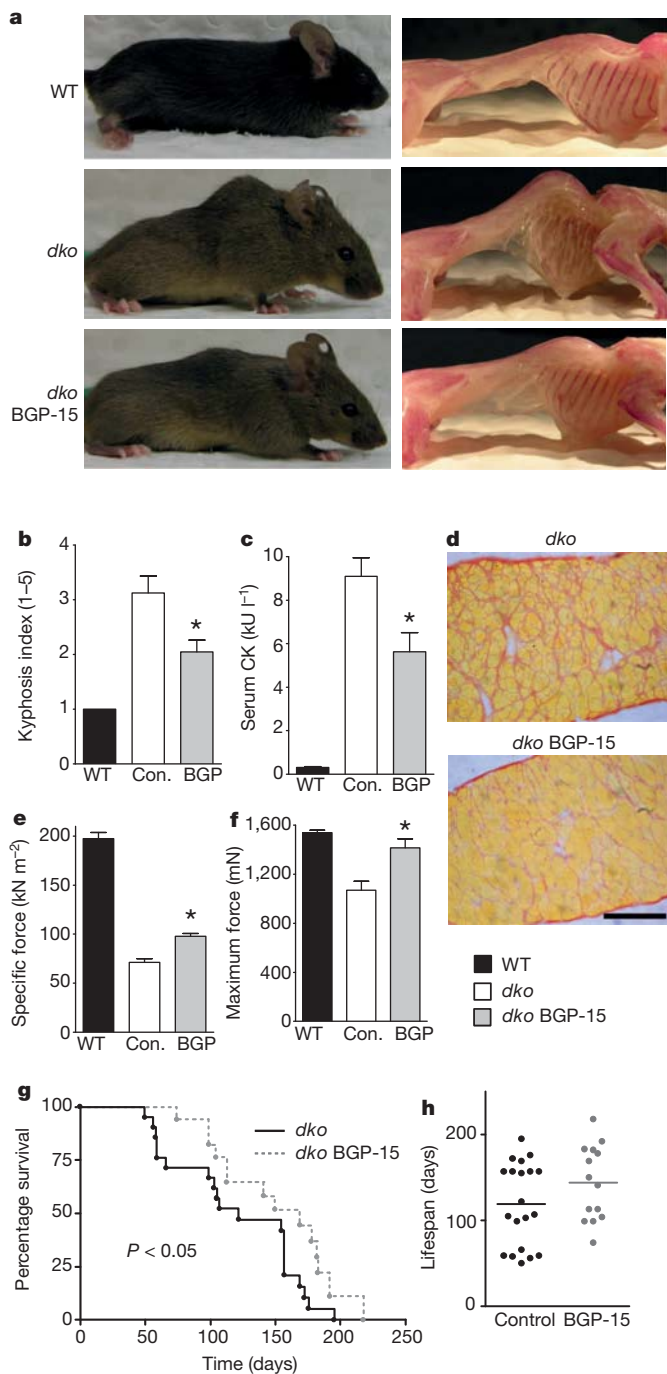


Figure 4 | Treatment with BGP-15 decreases kyphosis (spinal curvature), improves muscle function and prolongs lifespan in severely dystrophic *dko* mice. **a**, Left: representative photographs of WT (C57BL/10), *dko* and BGP-15-treated *dko* mice. Right: representative eviscerated skeletal preparations showing bone structure (pink), and highlighting spinal curvature in *dko* mice. **b**, Spinal curvature was quantified (1–5) in WT, *dko* and BGP-15-treated *dko* mice. Asterisk, $P < 0.05$, $n \geq 10$. **c**, Whole-body muscle breakdown, measured by serum CK levels. Asterisk, $P < 0.05$, $n \geq 8$. **d**, Representative images of collagen infiltration in diaphragm sections. **e**, Specific force of diaphragm muscle strips measured *in vitro*. Asterisk, $P < 0.01$, $n \geq 9$. **f**, Maximal force production in TA measured *in situ*. Asterisk, $P < 0.01$, $n \geq 9$. **g**, Survival curve of untreated (control) and BGP-15-treated *dko* mice. $P < 0.05$, $n \geq 14$. **h**, Scatter-plot of *dko* lifespan with a line showing median survival. Scale bar, 200 μ m. WT data are used as a reference control. Data are shown as means \pm s.e.m.

obstacles need to be overcome before these approaches can be considered safe and effective. Until these concerns are obviated, alternative (and potentially synergistic) therapies, such as pharmacological

induction of Hsp72, could delay the disease progression to allow many patients to benefit from perfected treatments.

In a recent clinical trial for patients with insulin resistance, treatment with BGP-15 (200 mg and 400 mg once a day for 28 days) significantly increased sensitivity to insulin. The dose schedule of BGP-15 was well tolerated, with both doses being safe: there were no clinically significant changes in physical status or in laboratory or electrocardiogram parameters²⁹. Given that BGP-15 is currently used in clinical trials for other pathologies, our findings identify it as a tangible and realistic treatment method for patients with DMD in the near future.

METHODS SUMMARY

Animals. All experiments were approved by the Animal Ethics Committee of The University of Melbourne and conducted in accordance with the Australian code of practice for the care and use of animals for scientific purposes as stipulated by the National Health and Medical Research Council (NHMRC, Australia). Male mice were used for all experiments. Wild-type (C57BL/10) and dystrophic *mdx* mice were sourced from the Animal Resources Centre (Canning Vale, Western Australia).

Muscle functional analysis. Mice were anaesthetized with sodium pentobarbitone (Nembutal) such that they were unresponsive to tactile stimuli. Contractile properties of diaphragm muscle strips were assessed *in vitro*.

Morphological analysis. Muscles were trimmed of tendons and adhering non-muscle tissue, mounted in embedding medium, frozen in liquid-nitrogen-cooled isopentane, and stored at -80°C . Transverse muscle sections were cryosectioned from the mid-belly of each muscle. Muscle collagen content was assessed from Van Gieson's stained cross-sections that were quantified.

SERCA activity assay. Ca^{2+} -dependent SERCA activity was assessed in isolated enriched SR vesicles and whole-muscle homogenates. For whole-muscle homogenates, muscles were surgically excised from anaesthetized mice and stored at -80°C for subsequent analyses. For enriched SR vesicles, mixed hindlimb muscles (quadriceps, gastrocnemius, extensor digitorum longus, soleus and plantaris) and diaphragm muscles were homogenized and subjected to sucrose gradient differential centrifugation using a Thermo Scientific Sorvall WX100 ultracentrifuge with a T-890 fixed-angle rotor. During the entire homogenization and SR vesicle isolation procedures, samples were immersed in ice to avoid temperature-dependent decreases in SERCA activity.

Skeletal preparation. To reveal skeletal architecture, mice (after death) were skinned, eviscerated and placed in a KOH solution (1.5% w/v), for 5 days. KOH solution was replaced and a small amount of alizarin red was added to stain calcium deposits; the preparation was left for a further 5 days.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 January; accepted 21 February 2012.

Published online 4 April 2012.

- Emery, A. E. The muscular dystrophies. *Lancet* **359**, 687–695 (2002).
- Ervasti, J. M. & Campbell, K. P. Membrane organization of the dystrophin-glycoprotein complex. *Cell* **66**, 1121–1131 (1991).
- Koenig, M., Monaco, A. P. & Kunkel, L. M. The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* **53**, 219–228 (1988).
- Blake, D. J., Weir, A., Newey, S. E. & Davies, K. E. Function and genetics of dystrophin and dystrophin-related proteins in muscle. *Physiol. Rev.* **82**, 291–329 (2002).
- Straub, V., Rafael, J. A., Chamberlain, J. S. & Campbell, K. P. Animal models for muscular dystrophy show different patterns of sarcolemmal disruption. *J. Cell Biol.* **139**, 375–385 (1997).
- Turner, P. R., Westwood, T., Regen, C. M. & Steinhardt, R. A. Increased protein degradation results from elevated free calcium levels found in muscle from *mdx* mice. *Nature* **335**, 735–738 (1988).
- Grady, R. M. *et al.* Skeletal and cardiac myopathies in mice lacking utrophin and dystrophin: a model for Duchenne muscular dystrophy. *Cell* **90**, 729–738 (1997).
- Deconinck, A. E. *et al.* Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. *Cell* **90**, 717–727 (1997).
- Gervasio, O. L., Whitehead, N. P., Yeung, E. W., Phillips, W. D. & Allen, D. G. TRPC1 binds to caveolin-3 and is regulated by Src kinase—role in Duchenne muscular dystrophy. *J. Cell Sci.* **121**, 2246–2255 (2008).
- Fong, P. Y., Turner, P. R., Denetclaw, W. F. & Steinhardt, R. A. Increased activity of calcium leak channels in myotubes of Duchenne human and *mdx* mouse origin. *Science* **250**, 673–676 (1990).
- Bellinger, A. M. *et al.* Hypernitrosylated ryanodine receptor calcium release channels are leaky in dystrophic muscle. *Nature Med.* **15**, 325–330 (2009).
- Nicolas-Metral, V., Raddatz, E., Kucera, P. & Ruegg, U. T. *Mdx* myotubes have normal excitability but show reduced contraction-relaxation dynamics. *J. Muscle Res. Cell Motil.* **22**, 69–75 (2001).

13. Tutdibi, O., Brinkmeier, H., Rudel, R. & Fohr, K. J. Increased calcium entry into dystrophin-deficient muscle fibres of MDX and ADR-MDX mice is reduced by ion channel blockers. *J. Physiol. (Lond.)* **515**, 859–868 (1999).
14. Goonasekera, S. A. *et al.* Mitigation of muscular dystrophy in mice by SERCA overexpression in skeletal muscle. *J. Clin. Invest.* **121**, 1044–1052 (2011).
15. Morine, K. J., Sleeper, M. M., Barton, E. R. & Sweeney, H. L. Overexpression of SERCA1a in the *mdx* diaphragm reduces susceptibility to contraction-induced damage. *Hum. Gene Ther.* **12**, 1735–1739 (2010).
16. Porter, J. D. *et al.* A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient *mdx* mice. *Hum. Mol. Genet.* **11**, 263–272 (2002).
17. Acharyya, S. *et al.* Interplay of IKK/NF- κ B signaling in macrophages and myofibres promotes muscle degeneration in Duchenne muscular dystrophy. *J. Clin. Invest.* **117**, 889–901 (2007).
18. Kolodziejczyk, S. M. *et al.* Activation of JNK1 contributes to dystrophic muscle pathogenesis. *Curr. Biol.* **11**, 1278–1282 (2001).
19. Monici, M. C., Aguenouz, M., Mazzeo, A., Messina, C. & Vita, G. Activation of nuclear factor- κ B in inflammatory myopathies and Duchenne muscular dystrophy. *Neurology* **60**, 993–997 (2003).
20. Senf, S. M., Dodd, S. L., McClung, J. M. & Judge, A. R. Hsp70 overexpression inhibits NF- κ B and Foxo3a transcriptional activities and prevents skeletal muscle atrophy. *FASEB J.* **22**, 3836–3845 (2008).
21. Park, H. S., Lee, J. S., Huh, S. H., Seo, J. S. & Choi, E. J. Hsp72 functions as a natural inhibitory protein of c-Jun N-terminal kinase. *EMBO J.* **20**, 446–456 (2001).
22. Chung, J. *et al.* HSP72 protects against obesity-induced insulin resistance. *Proc. Natl Acad. Sci. USA* **105**, 1739–1744 (2008).
23. Tupling, A. R. *et al.* HSP70 binds to the fast-twitch skeletal muscle sarco(endo)plasmic reticulum Ca^{2+} -ATPase (SERCA1a) and prevents thermal inactivation. *J. Biol. Chem.* **279**, 52382–52389 (2004).
24. Bornman, L., Polla, B. S., Lotz, B. P. & Gericke, G. S. Expression of heat-shock/stress proteins in Duchenne muscular dystrophy. *Muscle Nerve* **18**, 23–31 (1995).
25. Lou, J. S., Weiss, M. D. & Carter, G. T. Assessment and management of fatigue in neuromuscular disease. *Am. J. Hosp. Palliat. Care* **27**, 145–157 (2010).
26. Finsterer, J. Cardiopulmonary support in Duchenne muscular dystrophy. *Lung* **184**, 205–215 (2006).
27. Stedman, H. H. *et al.* The *mdx* mouse diaphragm reproduces the degenerative changes of Duchenne muscular dystrophy. *Nature* **352**, 536–539 (1991).
28. Briguët, A., Courdier-Fruh, I., Foster, M., Meier, T. & Magyar, J. P. Histological parameters for the quantitative assessment of muscular dystrophy in the *mdx*-mouse. *Neuromuscul. Disord.* **14**, 675–682 (2004).
29. Literati-Nagy, B. *et al.* Improvement of insulin sensitivity by a novel drug, BGP-15, in insulin-resistant patients: a proof of concept randomized double-blind clinical trial. *Horm. Metab. Res.* **41**, 374–380 (2009).
30. Morimoto, R. I. Cells in stress: transcriptional activation of heat shock genes. *Science* **259**, 1409–1410 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Koopman, J. G. Ryall and G. I. Lancaster for comments; J. Trieu, B. G. Gleeson, T. Naim and A. Chee for technical support; and C. Angelini and the Neuromuscular bank of tissues and DNA samples – Telethon Network of Genetic Biobanks for the provision of human muscle specimens. We thank N-Gene R&D Inc. USA for providing the BGP-15 compound. This study was supported in part by research grants from the National Health and Medical Research Council (NHMRC; project numbers 1009114 to G.S.L. and 472650 and 1004441 to M.A.F.), Association Française contre les Myopathies (France, to G.S.L.) and the Muscular Dystrophy Association (USA, to G.S.L.). M.A.F. is a Senior Principal Research Fellow of the NHMRC. A.P.R. was supported by a NHMRC Biomedical career Development Award. S.L. was supported by a postdoctoral fellowship from the Swiss National Science Foundation. S.M.G. was supported by a National Heart Foundation Postgraduate Scholarship (Australia). D.C.H. was supported by a National Heart Foundation Post-Doctoral Fellowship.

Author Contributions S.M.G., J.D.S., C.v.d.P., M.A.F. and G.S.L. conceived and designed the experiments. S.M.G., T.A.S., C.v.d.P., D.C.H. and J.E.C. performed the experiments. M.A.F. and K.E.D. facilitated experiments through the provision of mice and experimental compounds. A.P.R. and S.L. performed experiments on muscle samples from patients with DMD and from controls. S.M.G., D.C.H., M.A.F., C.v.d.P. and G.S.L. analysed the data. S.M.G. and G.S.L. wrote the manuscript. All authors checked for scientific content and contributed to the final drafting of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.S.L. (gsl@unimelb.edu.au).

METHODS

Animals. Female *mdx* mice were crossed with male mice expressing a rat inducible Hsp72 transgene under the control of a β -actin promoter³¹. F₁ generation males were mated with female *mdx* mice to yield an equal proportion of *mdx*^{TG(+)} and *mdx*^{TG(-)} littermate controls. WT^{TG(-)} and WT^{TG(+)} mice were generated as described previously²². Genotyping was performed by PCR using primers described previously³². C57BL/10 and *mdx* mice used for heat therapy and BGP-15 treatment studies were obtained from the Animal Resources Centre (Canning Vale, Western Australia). Dystrophic *mdx* mice were treated from 4 weeks to 9 weeks of age (5 weeks) or from 4 weeks to 16 weeks of age (12 weeks) with BGP-15 (15 mg kg⁻¹ per day, oral gavage; N-Gene R&D Inc.). BGP-15 is a pharmacological inducer of Hsp72, which has been shown to be safe and well tolerated in FDA-approved Phase II clinical trials for diabetes and insulin resistance^{29,33}. For heat therapy, the mice were anaesthetized and the core body temperature was raised to about 40 °C with an infrared thermometer for 30 min every fourth day; body temperature was monitored from the external auditory meatus as described previously³⁴. The *dko* mice were bred in the Biological Research Facility at the University of Melbourne⁸. The *dko* mice are utrophin-null on an *mdx* background and show a severe and progressive muscular dystrophy that more closely mimics the phenotypic characteristics of DMD^{7,8}. The *dko* mice were treated with BGP-15 from 3–4 weeks to 10 weeks of age. For the survival study, *dko* mice were treated from 3–4 weeks until death or humane killing in accordance with the Animal Ethics Committee of The University of Melbourne (AEC). All experiments were approved by the AEC and conducted in accordance with the Australian code of practice for the care and use of animals for scientific purposes as stipulated by the National Health and Medical Research Council (Australia). Male mice were used for all experiments.

Muscle functional analyses. Mice were anaesthetized with sodium pentobarbitone (Nembutal; 60 mg kg⁻¹ intraperitoneal; Sigma-Aldrich, New South Wales, Australia) such that they were unresponsive to tactile stimuli. Contractile properties of diaphragm muscle strips were assessed *in vitro* as described in detail previously³⁷. In brief, strips of diaphragm muscle were bathed in oxygenated Krebs solution at 25 °C in a custom-made organ bath. Preparations were stimulated by supramaximal 0.2-ms square-wave pulses of 450 ms duration delivered by means of platinum plate electrodes flanking both sides of the muscle. Contractile properties of the TA muscle were assessed *in situ*³⁵. In brief, TA muscles maintained at 37 °C were stimulated by supramaximal (14 V) 0.2-ms square-wave pulses of 350 ms duration delivered by means of two wire electrodes resting on the peroneal nerve. All muscles were adjusted to optimum muscle length (*L*₀), determined from maximum isometric twitch force (*P*_i). Maximum isometric tetanic force (*P*₀) was recorded from the plateau of the frequency–force relationship, and normalized to muscle cross-sectional area (specific force; *sP*₀), for comparisons between groups where appropriate^{36–38}. The susceptibility of TA and diaphragm muscles to contraction-induced injury was determined from the protocols that we have described in detail previously^{36–38}. Isolated muscles were maximally activated to produce isometric force and then stretched to perform an eccentric contraction (at a velocity of 2*L*_fs⁻¹) at progressively increasing magnitudes of stretch. Maximum isometric force was determined before each eccentric contraction³⁷.

Morphological analysis. Muscles were trimmed of tendons and adhering non-muscle tissue, mounted in embedding medium, frozen in liquid-nitrogen-cooled isopentane, and stored at -80 °C. Transverse muscle sections (5 μ m) were cryosectioned from the mid-belly of each muscle. Sections were stained with haematoxylin/eosin to reveal general muscle architecture. Cross-sectional area (CSA) and minimal Feret's diameter was assessed in about 500 fibres from each section of diaphragm muscle, with the use of Carl Zeiss software (Axiovision 4.6.3). The minimal Feret's diameter is the minimum distance between opposing parallel tangents of a muscle fibre³⁹, and the variance coefficient is a highly sensitive parameter used for detecting differences between dystrophic and otherwise healthy muscles³⁹. Muscle collagen content was assessed from Van Gieson's stained cross-sections that were quantified with Axiovision 4.6.3 software.

Enriched SR isolation. Mixed hindlimb muscles (quadriceps, gastrocnemius, extensor digitorum longus, soleus and plantaris) and diaphragm muscles were diluted in ice-cold homogenization buffer (250 mM sucrose, 5 mM HEPES pH 7.0, 0.2% Na₂S₂O₈). Protease inhibitor cocktail (P-8340; Sigma-Aldrich, Castle Hill, New South Wales, Australia) was added immediately before homogenization at 5 μ l per 100 mg muscle wet weight. The muscles were minced on ice with surgical scissors and homogenized with a Polytron homogenizer (PT2100; Kinematica, Inc. Dispersing and Mixing Technology, New York) at a power setting of 21 for three 20-s bursts separated by 45-s breaks on ice. To obtain a purified and enriched SR membrane fraction, the homogenates were centrifuged at 5,500g for 10 min at 4 °C to remove insoluble material. The supernatants were harvested and centrifuged at 12,500g for 18 min at 4 °C. The pellet was discarded and the supernatant was

centrifuged for a second time at 12,500g for 18 min at 4 °C. Supernatants were transferred to ultracentrifuge tubes, which were balanced and centrifuged at 50,000g (24,200 r.p.m. on a T-890 fixed-angle rotor; Thermo Scientific Sorvall WX100 ultracentrifuge) for 1 h at 4 °C. Supernatants were discarded and pellets were resuspended in 5 ml of homogenization buffer containing 600 mM KCl and incubated on ice for 30 min. Samples were centrifuged at 15,000g for 10 min at 4 °C to pellet and remove mitochondria. Supernatants were centrifuged again at 50,000g for 1 h at 4 °C. The final pellet (enriched SR membrane vesicles) was resuspended in homogenization buffer. Protein content was determined in triplicate⁴⁰.

SERCA activity assay. Ca²⁺-dependent SERCA activity was assessed in isolated enriched SR vesicles and whole-muscle homogenates on the basis of the methods of Leberer and colleagues, as described previously⁴¹. For whole-muscle homogenates, muscles were surgically excised from anaesthetized mice and stored at -80 °C. Muscle samples (about 20–50 mg) were diluted in about 200 μ l of ice-cold homogenization buffer. Protease inhibitor cocktail (P-8340) was added immediately before use at a concentration of 5 μ l per 100 mg of muscle tissue. Muscles were homogenized with a hand-held glass homogenizer, and then centrifuged for 10 min at 5,500g at 4 °C. Supernatant protein concentration was determined in triplicate⁴⁰. Protein concentration was adjusted to 10 mg ml⁻¹, when possible, with homogenization buffer. SERCA activity in whole-muscle homogenates was determined in reaction buffer (200 mM KCl, 20 mM HEPES, pH 7.0, 15 mM MgCl₂, 10 mM Na₂S₂O₈, 10 mM phosphoenolpyruvate, 5 mM ATP, 1 mM EGTA). SERCA activity in enriched SR vesicles was determined in reaction buffer (100 mM KCl, 20 mM HEPES, 10 mM MgCl₂, 10 mM Na₂S₂O₈, 10 mM phosphoenolpyruvate, 5 mM ATP, and 1 mM EGTA). The pH of both reaction buffers was adjusted to 7.0 at 37 °C. Immediately before starting the reaction, 18 U ml⁻¹ PK, 18 U ml⁻¹ lactate dehydrogenase (LDH), 5 μ l NADH (100 mM), 1 μ M calyculin A-23187 (Sigma-Aldrich) and about 10 μ l of whole-muscle homogenate or about 3 μ l of enriched SR vesicles were added to 1 ml of reaction buffer in a plastic cuvette. Cuvettes were loaded into a spectrophotometer and A₃₄₀ was measured at 37 °C (Multiscan Spectrum; Thermo Electron, Waltham, Massachusetts). Maximal (*V*_{max}) and Ca²⁺-dependent SERCA activities were determined by progressively adding Ca²⁺ until a plateau or maximal activity was reached. The specific SERCA inhibitor 2',5'-di(tert-butyl)-1,4-benzohydroquinone (TBQ) was added to a final concentration of 40 μ M to determine basal activity. SERCA enzyme kinetic parameters, determined from a regression analysis, were the Ca²⁺ concentration required to elicit 50% maximal SERCA activity ([Ca²⁺]₅₀) and the Hill coefficient, which is a measure of the cooperativity of Ca²⁺ binding of the SERCA enzyme. SERCA activity was graphed against added Ca²⁺ concentration. Non-regression analysis was performed using the following variable-slope sigmoidal dose–response relationship, using Graphpad Prism v.3.02 (GraphPad Software Inc., San Diego):

$$Y = Y_{\text{bottom}} + Y_{\text{top}} - \frac{Y_{\text{bottom}}}{1 + 10^{\log[\text{Ca}]_{50} - X/n_H}}$$

Peroxyntirite-mediated SERCA inhibition assay. SERCA activity in enriched SR vesicles after incubation in various concentrations of peroxyntirite ([ONOO⁻]) was measured. [ONOO⁻] was measured spectrophotometrically (ϵ = 1,670 M cm⁻¹ at 302 nm, about 68 mM in stock). Enriched SR vesicle samples were added to reaction buffer (as described previously) in the presence of ONOO⁻ (0, 10, 50, 100, 250, 500 or 1,000 μ M) and incubated for 5 min at 37 °C in a plastic cuvette. ONOO⁻ and byproducts were quenched by the addition of 1 mM dithiothreitol (to prevent the inactivation of PK and LDH subsequently added); 18 U ml⁻¹ PK, 18 U ml⁻¹ LDH and 5 μ l of NADH (100 mM) were then added. Cuvettes were loaded into a spectrophotometer, and A₃₄₀ was measured at 37 °C. Maximal SERCA activity was determined by adding Ca²⁺ until a plateau or maximal activity was reached. Once maximal activity had been determined, the specific SERCA inhibitor TBQ was added to a final concentration of 40 μ M to determine basal activity. Maximal SERCA activity in each sample was assessed independently at various ONOO⁻ concentrations (10, 50, 100, 250, 500 or 1,000 μ M) and expressed as a percentage of the maximal SERCA activity in the absence of ONOO⁻.

Human DMD samples. Human muscle specimens were sourced from the Telethon Network of Genetic Biobanks, Italy. Biopsies were taken from the vastus lateralis muscle of patients with DMD (aged 1–9 years) or healthy controls (aged 18–25 years) using either a Bergström or open biopsy technique. All biopsies were frozen immediately in liquid nitrogen and stored at -80 °C until analysis. Muscle samples were homogenized in RIPA buffer (Millipore, Billerica, Massachusetts) and rotated for 1 h at 4 °C, followed by centrifugation for 15 min at 4 °C. The supernatant was collected and protein concentration was determined by means of the bicinchoninic acid protein assay (Pierce Biotechnology). Electrophoresis was performed by 10% SDS–PAGE in a buffer containing 12 mM Tris–HCl pH 8.8,

200 mM glycine and 0.1% SDS. Protein transfer was performed in a buffer containing 12 mM Tris-HCl pH 8.3, 200 mM glycine and 10% methanol with the use of poly(vinylidene difluoride) membranes. The membranes were blocked with 5% BSA in PBS, after which they were incubated overnight at 4 °C with the primary antibody against Hsp70 (SPA-812; Stressgen) diluted 1:1000 in PBS. After being washed, the membranes were incubated for 1 h with a goat anti-rabbit IgG antibody labelled with an infrared-fluorescent 800-nm dye (Alexa Fluor 800; Invitrogen) diluted 1:5,000 in PBS containing 50% Odyssey blocking buffer (LI-COR Biosciences) and 0.01% SDS. After being washed, the proteins were exposed on an Odyssey Infrared Imaging System (LI-COR Biosciences) and individual protein band optical densities were determined with ImageJ Software (National Institutes of Health). The blots were normalized against glyceraldehyde-3-phosphate dehydrogenase (GAPDH) protein (G8795; Sigma-Aldrich, Sydney, Australia).

Evans blue dye uptake. To quantify muscle damage and areas of focal necrosis, EBD (1% w/v) was injected intraperitoneally (10 µl per gram body mass). Muscles were frozen 20 h after the EBD injection. Sections 10 µm thick were cut on a cryostat, and EBD was detected as red autofluorescence with the use of a fluorescence microscope.

Wire hang test. A wire hang test was employed to assess whole-body muscle strength and endurance. Mice were placed on a wire mesh grid, which they gripped; the grid was then inverted. Latency-to-fall onto a padded surface was recorded in three successive trials, with the average of the best two trials used for analyses.

Single muscle fibre analysis. Mice were anaesthetized with sodium pentobarbitone (Nembutal) such that they were unresponsive to tactile stimuli, and the extensor digitorum longus and/or soleus muscle was surgically excised for single fibre analysis. The muscle was blotted on filter paper and placed in a Petri dish containing paraffin oil at room temperature (23 ± 2 °C). Muscles were pinned at resting length to the base of a dish. Single muscle fibres were isolated from as close to the surface of the muscle as possible, and fine forceps were used to peel the sarcolemma away from the contractile apparatus under a dissecting microscope. The mechanically skinned fibre was then attached to one end of a piezoresistive force transducer (AE801 SensoNor; Horten) using braided silk (size 10, 0.2 mm; Deknatel), and the other end of the fibre was clamped between a pair of forceps fixed to a micromanipulator. All experiments were conducted at room temperature. All solutions had pH of 7.10 ± 0.01, and the free Mg²⁺ concentration ([Mg²⁺]) was 1 mM, unless specified otherwise. Free [Ca²⁺] at 0.1 µM or more was verified with a Ca²⁺-sensitive electrode (Orion Research).

Caffeine-induced force responses and SR Ca²⁺ accumulation. Initially, mechanically skinned muscle fibres were equilibrated for 30 s in a wash solution followed by thorough depletion of SR Ca²⁺ stores, achieved by transferring the fibre preparation into a release solution containing 30 mM caffeine and 0.02 mM free Mg²⁺. The presence of 0.5 mM EGTA in the release solution ensured that the level of Ca²⁺ during caffeine-induced release did not maximally activate the contractile apparatus, which is necessary to allow quantitative evaluation of the amount of Ca²⁺ released. Ca²⁺ release from the SR was estimated from the relative areas under the caffeine-induced force response. The fibre was left in the release solution for 2 min to ensure complete SR Ca²⁺ depletion, before being washed for 30 s. Thereafter, the SR was reloaded with Ca²⁺ in load solution (0.2 µM Ca²⁺ (pCa 6.7), where pCa = -log₁₀[Ca²⁺]) for increasing durations (10, 20, 30 and 60 s), before being equilibrated for 30 s in wash solution; subsequently, SR Ca²⁺ was released in release solution. Data were fitted with a standard exponential association equation to give the rate at which the SR accumulated Ca²⁺ (in s⁻¹) but not the amount of SR Ca²⁺ accumulated. SR accumulation is an indirect measure of SERCA activity⁴².

SR Ca²⁺ leak. The percentage of Ca²⁺ lost from the SR as a result of the passive leak was also assessed. The fibre was loaded for 20 s in loading solution. The fibre preparation was then placed in wash solution for 30 s followed by SR Ca²⁺ content released in release solution (Ca²⁺ leak in 30 s). The fibre preparation was placed in wash solution before reloading for 20 s in load solution and transferred to wash solution for 90 s; the remaining SR Ca²⁺ was released in release solution (Ca²⁺ leak in 90 s). The 30-s Ca²⁺ leak was then repeated, and the area (corrected for proportionality between area and SR Ca²⁺ content) under the test run was divided by the average of the areas under the caffeine-induced force responses in the controls before and after the test run. This gave an estimate of the fraction of SR Ca²⁺ leaked over a 60-s leak period.

Relative SR Ca²⁺ sensitivity. To determine the effect of Hsp72 overexpression on the ryanodine receptor (RyR), a caffeine dose-response curve was determined from the forces produced by the contractile apparatus after SR Ca²⁺ release induced by low caffeine concentrations. Each fibre was prepared by completely depleting the SR of Ca²⁺ with 30 mM caffeine followed by a 30-s SR Ca²⁺ reloading duration. The peak force of caffeine-induced contraction was determined in a

series of potassium hexamethylene-diamine tetraacetic acid solutions containing 2, 3, 5 and 7 mM caffeine and 50 µM EGTA, with complete SR Ca²⁺ depletion and 30-s reloading duration of the SR with Ca²⁺ between each caffeine contraction. The peaks of the caffeine-induced contractions were then normalized as a percentage of maximum Ca²⁺-activated force to estimate the sensitivity of the RyR to caffeine. This determined whether overexpression of Hsp72 directly affected the function of the RyR.

Properties of the contractile apparatus. After SR properties were investigated, the single muscle fibres were equilibrated in a relaxing solution (pCa > 9) for 2 min. Fibres were placed in a maximum Ca²⁺-activating solution (pCa ≈ 4.5) until force reached the maximal value (maximum Ca²⁺-activated force) and then placed back in the relaxing solution for a further 2 min. Force responses were generated by exposing the fibre to activating solutions of progressively lower pCa (higher [Ca²⁺]) in a stepwise fashion. The force response generated at each pCa was expressed as a percentage of the interpolated values for maximum Ca²⁺-activated force. Data points were fitted with a Hill equation producing two parameters: the pCa₅₀ (the pCa producing half-maximum force) and n_H (the Hill coefficient, indicative of the steepness of the force-pCa relationship).

Western blotting. After the determination of SERCA activity in whole-muscle homogenates and enriched SR vesicles, electrophoresis was performed on the same samples for the quantification of protein expression. Equal amounts of protein were resolved in SDS buffer, heated for 5 min at 95 °C and separated on SDS-polyacrylamide gels. Separated proteins were transferred to poly(vinylidene difluoride) membranes (0.45-µm Immobilon-P; Millipore). Membranes were blocked with 5% non-fat milk powder (or BSA) in Tris-buffered saline containing Tween 20 for 1 h and incubated overnight at 4 °C with appropriate antibody dilutions. Antibodies against SERCA 1 (Affinity Bioreagents), SERCA 2a (Affinity Bioreagents), Hsp72 (Stressgen), α -tubulin (ECM-Biosciences) and GAPDH (Sigma) were used. Membranes were digitized with a chemiluminescent detection and imaging system (ChemiDoc XRS; Bio-Rad Laboratories, Hercules, California, USA) and band densities were quantified with Quantity One analysis software (version 4.6.1; Bio-Rad Laboratories).

Real-time PCR. Real-time PCR was performed as described previously⁴³. Each sample was run in triplicate. The mean reading of each triplicate was converted to an absolute content by using a DNA standard curve, based on a serial dilution of DNA standard (100–10,000 ng ml⁻¹; Molecular Probes) run together with the samples on each plate. Gene expression was quantified by normalizing the logarithmic cycle threshold (C_t) value (2C_t) to the cDNA content of each sample to obtain the expression 2C_t/cDNA content (ng ml⁻¹). In Supplementary Fig. 5b, mRNA for SERCA 1 and SERCA 2a was normalized to the eukaryotic 18S house-keeping gene.

Creatine kinase analysis. Serum CK activity was analysed as an overall measure of whole-body muscle breakdown. Blood was collected from the tail vein and centrifuged for 10 min at 10,000g and 4 °C to isolate the serum fraction. Serum CK activity was then determined with a commercially available creatine kinase assay kit (ECPK-100) in accordance with the manufacturer's instructions (BioAssay Systems).

Skeletal preparation. Skeletal architecture was revealed as described previously⁴⁴. After death, the mice were skinned, eviscerated and placed in a KOH solution (1.5% w/v), for 5 days. KOH solution was replaced and a small amount of alizarin red was added to stain calcium deposits; the preparation was left for a further 5 days to produce the specimens shown in Fig. 4a.

Statistical analysis. All values are presented as means ± s.e.m. Unpaired Student's *t*-tests were used to compare between two groups. When comparisons were being made between more than two groups, a one-way analysis of variance was used with Newman-Keuls post-hoc multiple comparison test.

- Marber, M. S. *et al.* Overexpression of the rat inducible 70-kD heat stress protein in a transgenic mouse increases the resistance of the heart to ischemic injury. *J. Clin. Invest.* **95**, 1446–1456 (1995).
- Taleb, M. *et al.* Hsp70 inhibits aminoglycoside-induced hair cell death and is necessary for the protective effect of heat shock. *J. Assoc. Res. Otolaryngol.* **9**, 277–289 (2008).
- Literati-Nagy, B. *et al.* Beneficial effect of the insulin sensitizer (HSP inducer) BGP-15 on olanzapine-induced metabolic disorders. *Brain Res. Bull.* **20**, 340–344 (2010).
- Saegusa, Y. & Tabata, H. Usefulness of infrared thermometry in determining body temperature in mice. *J. Vet. Med. Sci.* **65**, 1365–1367 (2003).
- Gregorevic, P., Plant, D. R., Leeding, K. S., Bach, L. A. & Lynch, G. S. Improved contractile function of the *mdx* dystrophic mouse diaphragm muscle after insulin-like growth factor-I administration. *Am. J. Pathol.* **161**, 2263–2272 (2002).
- Schertzer, J. D., Ryall, J. G. & Lynch, G. S. Systemic administration of IGF-I enhances oxidative status and reduces contraction-induced injury in skeletal muscles of *mdx* dystrophic mice. *Am. J. Physiol.* **291**, 499–505 (2006).
- Schertzer, J. D., Gehrig, S. M., Ryall, J. G. & Lynch, G. S. Modulation of insulin-like growth factor (IGF)-I and IGF-binding protein interactions enhances skeletal

- muscle regeneration and ameliorates the dystrophic pathology in *mdx* mice. *Am. J. Pathol.* **171**, 1180–1188 (2007).
38. Gehrig, S. M., Koopman, R., Naim, T., Tjoakarfa, C. & Lynch, G. S. Making fast-twitch dystrophic muscles bigger protects them from contraction injury and attenuates the dystrophic pathology. *Am. J. Pathol.* **176**, 29–33 (2010).
39. Brigue, A., Courdier-Fruh, I., Foster, M., Meier, T. & Magyar, J. P. Histological parameters for the quantitative assessment of muscular dystrophy in the *mdx*-mouse. *Neuromuscul. Disord.* **14**, 675–682 (2004).
40. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
41. Schertzer, J. D. *et al.* β_2 -Agonist administration increases sarcoplasmic reticulum Ca^{2+} -ATPase activity in aged rat skeletal muscle. *Am. J. Physiol.* **288**, 526–533 (2005).
42. Plant, D. R. & Lynch, G. S. Depolarization-induced contraction and SR function in mechanically skinned muscle fibres from dystrophic *mdx* mice. *Am. J. Physiol.* **285**, C522–C528 (2003).
43. Murphy, K. T. *et al.* Antibody-directed myostatin inhibition in 21-mo-old mice reveals novel roles for myostatin signaling in skeletal muscle structure and function. *FASEB J.* **24**, 4433–4442 (2010).
44. Green, M. C. A rapid method for clearing and staining specimens for the demonstration of bone. *Ohio J. Sci.* **52**, 31–33 (1952).

Inositol-1,4,5-trisphosphate receptor regulates hepatic gluconeogenesis in fasting and diabetes

Yiguo Wang¹, Gang Li², Jason Goode¹, Jose C. Paz¹, Kunfu Ouyang³, Robert Screaton^{4,5,6}, Wolfgang H. Fischer¹, Ju Chen³, Ira Tabas^{2,7,8} & Marc Montminy¹

In the fasted state, increases in circulating glucagon promote hepatic glucose production through induction of the gluconeogenic program. Triggering of the cyclic AMP pathway increases gluconeogenic gene expression via the de-phosphorylation of the CREB co-activator CRTC2 (ref. 1). Glucagon promotes CRTC2 dephosphorylation in part through the protein kinase A (PKA)-mediated inhibition of the CRTC2 kinase SIK2. A number of Ser/Thr phosphatases seem to be capable of dephosphorylating CRTC2 (refs 2, 3), but the mechanisms by which hormonal cues regulate these enzymes remain unclear. Here we show in mice that glucagon stimulates CRTC2 dephosphorylation in hepatocytes by mobilizing intracellular calcium stores and activating the calcium/calmodulin-dependent Ser/Thr-phosphatase calcineurin (also known as PP3CA). Glucagon increased cytosolic calcium concentration through the PKA-mediated phosphorylation of inositol-1,4,5-trisphosphate receptors (InsP₃Rs), which associate with CRTC2. After their activation, InsP₃Rs enhanced gluconeogenic gene expression by promoting the calcineurin-mediated dephosphorylation of CRTC2. During feeding, increases in insulin signalling reduced CRTC2 activity via the AKT-mediated inactivation of InsP₃Rs. InsP₃R activity was increased in diabetes, leading to upregulation of the gluconeogenic program. As hepatic downregulation of InsP₃Rs and calcineurin improved circulating glucose levels in insulin resistance, these results demonstrate how interactions between cAMP and calcium pathways at the level of the InsP₃R modulate hepatic glucose production under fasting conditions and in diabetes.

We tested a series of Ser/Thr protein phosphatase inhibitors for their ability to block CRTC2 activation in response to glucagon. Exposure to the calcineurin inhibitor cyclosporine A (CsA) disrupted the glucagon-induced dephosphorylation and nuclear translocation of CRTC2, but okadaic acid, an inhibitor of PP1, PP2A and PP4 did not (Fig. 1a and Supplementary Fig. 1a). CsA and other calcineurin inhibitors also reduced cAMP response element (CRE)-luciferase (Luc) reporter activity (Fig. 1a and Supplementary Fig. 1b), but they had no effect in cells expressing phosphorylation-defective (Ser 171, 275 Ala) and therefore active forms of CRTC2 (Supplementary Fig. 1c–e).

On the basis of the ability of CsA to interfere with CRTC2 activation, we considered that calcineurin may promote the dephosphorylation of CRTC2 in response to glucagon. Supporting this idea, CRTC2 contains two consensus (PXIXIT) motifs that mediate an association with calcineurin^{3,4} (Supplementary Fig. 2a, b). Moreover, mutation of both motifs disrupted the glucagon-dependent dephosphorylation of CRTC2 (Fig. 1b) and prevented its nuclear translocation (Supplementary Fig. 2c), thereby down-regulating CRE-Luc activation (Fig. 1b).

On the basis of these results, we tested whether calcineurin modulates expression of the gluconeogenic program. Adenoviral over-expression of the calcineurin catalytic subunit in hepatocytes augmented CRTC2 dephosphorylation, CRE-Luc activity, and glucose secretion in

response to glucagon, whereas calcineurin knockdown had the opposite effect (Fig. 1c). Although calcineurin could, in principle, modulate CRTC2 activity indirectly through effects on cAMP signalling, calcineurin overexpression or knockdown did not alter the phosphorylation of cellular PKA substrates in cells exposed to glucagon (Supplementary Fig. 2d).

We examined next whether calcineurin modulates hepatic gluconeogenesis *in vivo*. Modest (twofold) overexpression of calcineurin in liver increased gluconeogenic gene expression, hepatic CRE-Luc activity, and fasting blood glucose concentrations (Fig. 1d and Supplementary Fig. 3a). Conversely, knockdown of hepatic calcineurin reduced expression of the gluconeogenic program and lowered circulating glucose levels (Fig. 1d and Supplementary Fig. 3b), demonstrating that this phosphatase contributes to fasting adaptation in the liver. Calcineurin seemed to stimulate gluconeogenesis via the CREB pathway; depletion of CRTC2 blocked the effects of calcineurin overexpression (Supplementary Fig. 4).

Realizing that calcineurin activity is dependent on increases in intracellular calcium, we tested whether the cAMP pathway stimulates calcium mobilization. Exposure of primary hepatocytes to glucagon triggered a rapid increase in cellular free calcium (Fig. 2a and Supplementary Fig. 5a); these effects were partially reversed by co-treatment with the PKA inhibitor H89 (Supplementary Fig. 5b). The rise in intracellular calcium seems to be critical for CRTC2 activation, because co-incubation with the calcium chelator BAPTA disrupted CRTC2 dephosphorylation and CRE-Luc activation in response to glucagon (Fig. 2b). Arguing against an effect of calcium on cAMP signalling, exposure to BAPTA did not block the PKA-mediated phosphorylation of CREB in response to glucagon.

We imagined that cAMP may increase calcium mobilization through the PKA-dependent phosphorylation of an intracellular calcium channel. In mass spectrometry studies to identify proteins that undergo phosphorylation by PKA in response to glucagon, we recovered the inositol 1,4,5-trisphosphate receptor 1 (InsP₃R1) from immunoprecipitates of phospho-PKA substrate antiserum (Supplementary Fig. 5c). InsP₃R1 and its related family members (InsP₃R2, InsP₃R3) are calcium release channels that promote the mobilization of endoplasmic reticulum calcium stores following their activation in response to extracellular signals^{5–9}. Moreover, cAMP agonists have also been shown to enhance InsP₃R receptor activity through PKA-mediated phosphorylation.

Inhibiting InsP₃Rs, either by exposure of hepatocytes to xestospongin C or by knockdown of all three InsP₃Rs, disrupted cytosolic calcium mobilization and calcineurin activation in response to glucagon and forskolin (Fig. 2a and Supplementary Fig. 6a). Moreover, xestospongin C treatment and InsP₃R knockdown also blocked the effects of glucagon on CRTC2 dephosphorylation, CRE-Luc activation, and induction of the gluconeogenic program (Fig. 2c and Supplementary Fig. 6a, b). We

¹Clayton Foundation Laboratories for Peptide Biology, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. ²Department of Medicine, Columbia University, New York, New York 10032, USA. ³Department of Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA. ⁴Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario K1H 8L1, Canada. ⁵Department of Pediatrics, University of Ottawa, Ottawa, Ontario K1H 8L1, Canada. ⁶Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario K1H 8L1, Canada. ⁷Department of Physiology and Cellular Biophysics, Columbia University, New York, New York 10032, USA. ⁸Department of Pathology and Cell Biology, Columbia University, New York, New York 10032, USA.

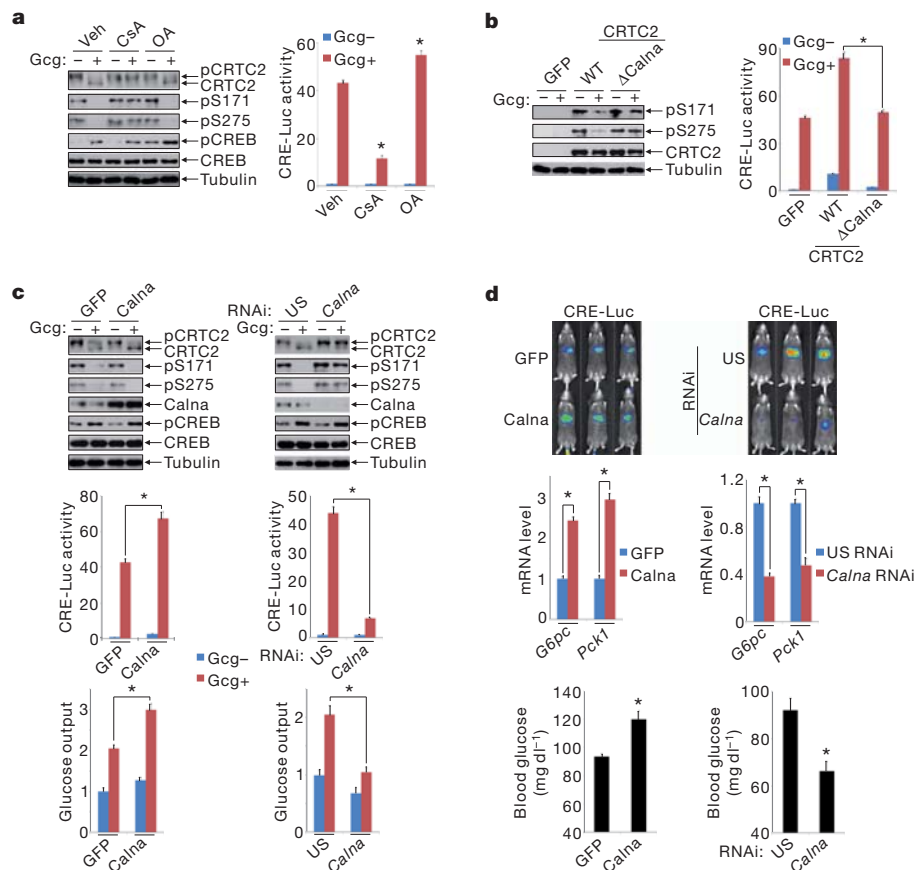


Figure 1 | Calcein promotes CRTC2 activation during fasting. **a**, Effect of Ser/Thr phosphatase inhibitors (okadaic acid (OA), CsA) on CRTC2 dephosphorylation and CRE-Luc reporter activation ($*P < 0.001$; $n = 3$). Gcg, glucagon; Veh, vehicle. **b**, Effect of glucagon on dephosphorylation (left) and activity (right) of wild-type (WT) and calcineurin-defective (Δ Calna) CRTC2 in hepatocytes ($*P < 0.001$; $n = 3$). **c**, Effect of calcineurin A overexpression

(left) or knockdown (right) on CRTC2 dephosphorylation (top), CRE-Luc reporter activity (middle, $*P < 0.001$; $n = 3$), and glucose output (bottom, $*P < 0.001$; $n = 3$) from hepatocytes. US, unspecific. **d**, Effect of hepatic calcineurin overexpression (left) or knockdown (right) on CRE-Luc activity, gluconeogenic gene (*Pck1*, *G6pc*) expression, and blood glucose concentrations in 6–8 h fasted mice ($*P < 0.01$; $n = 5$). Data are shown as mean \pm s.e.m.

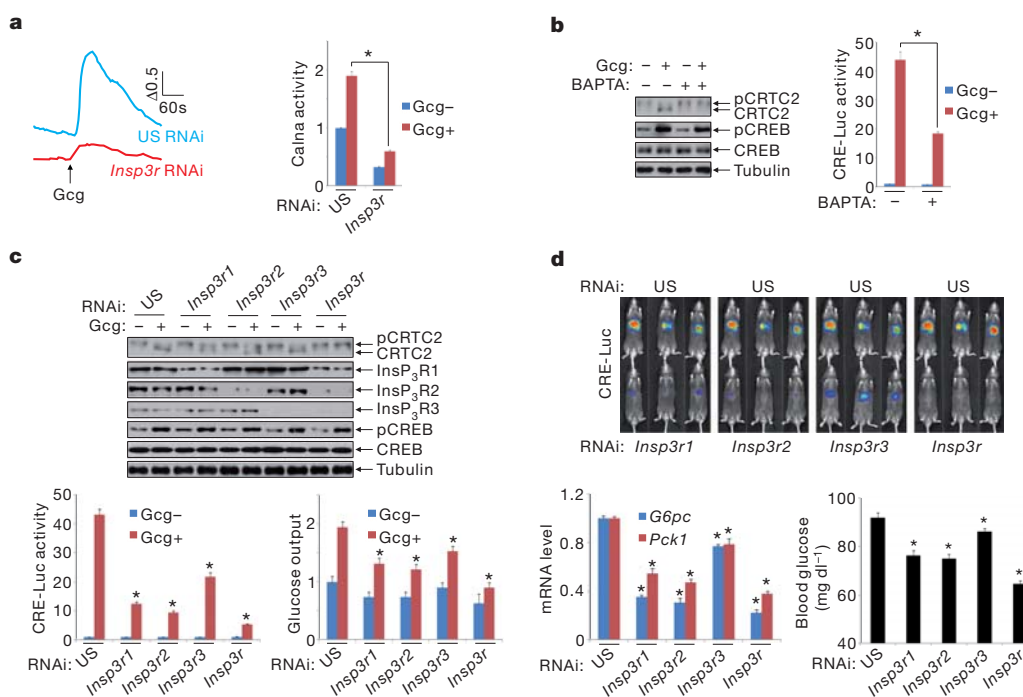


Figure 2 | Glucagon stimulates CRTC2 dephosphorylation via activation of InsP_3Rs . **a**, Effect of glucagon (Gcg) on calcium mobilization in hepatocytes by fluorescence imaging. Calcium mobilization and calcineurin activation following knockdown of all three InsP_3R family members shown ($*P < 0.001$; $n = 3$). **b**, Effect of calcium chelator (BAPTA) on CRTC2 dephosphorylation and CRE-Luc activation ($*P < 0.001$; $n = 3$). **c**, Effect of InsP_3R depletion on CRTC2 dephosphorylation, CRE-Luc activity, and glucose output from hepatocytes ($*P < 0.001$; $n = 3$). **d**, Effect of hepatic InsP_3R knockdown on CRE-Luc activity, blood glucose, and gluconeogenic gene expression ($*P < 0.01$; $n = 5$). Data are shown as mean \pm s.e.m.

confirmed the effects of InsP₃R depletion using hepatocytes from mice with a knockout of the InsP₃R2 (ref. 10), the predominant InsP₃R isoform in these cells (Supplementary Fig. 6c–e).

On the basis of these results, InsP₃Rs would also be expected to modulate fasting glucose production *in vivo*. Decreasing hepatic InsP₃R expression, either by knockdown of all three InsP₃Rs in liver or by targeted disruption of the *InsP3r2* gene, reduced fasting CRE-Luc activity, gluconeogenic gene expression, and circulating glucose concentrations, demonstrating the importance of these receptors in glucose homeostasis (Fig. 2d and Supplementary Fig. 7).

We tested whether glucagon modulates InsP₃R activity through PKA-mediated phosphorylation. Exposure of hepatocytes to glucagon increased the phosphorylation of InsP₃R1 as well as InsP₃R2 and InsP₃R3 by immunoblot assay with phospho-PKA substrate antiserum; these effects were blocked by the PKA inhibitor H89 (Fig. 3a and Supplementary Fig. 8a). Moreover, mutation of serine residues at consensus PKA sites in InsP₃R1 (Ser 1589, Ser 1756) to alanine completely disrupted InsP₃R1 phosphorylation in response to glucagon (Fig. 3b). As a result, overexpression of PKA-defective (S1589, S1756A) InsP₃R1 interfered with calcium mobilization and calcineurin activation, and it reduced CRE-Luc activation and glucose secretion from hepatocytes exposed to glucagon (Fig. 3b–d).

Similar to glucagon, fasting also stimulated hepatic InsP₃R1 phosphorylation at Ser 1589 and Ser 1756 (Supplementary Fig. 8b). And overexpression of PKA-defective InsP₃R1 reduced fasting CRE-Luc induction, calcineurin activation, and gluconeogenic gene expression, leading to lower circulating glucose concentrations (Fig. 3e and Supplementary Fig. 8c, d). Taken together, these results support an important role for the PKA-mediated phosphorylation of InsP₃R in hepatic gluconeogenesis.

We considered that the proximity of CRTC2 to the calcium signaling machinery may be important for its activation. Supporting this

notion, CRTC2 was found to associate with InsP₃R1 via its amino-terminal CREB binding domain (CBD) in co-immunoprecipitation assays (Fig. 3f and Supplementary Fig. 9a–d). Moreover, CRTC2 was enriched in endoplasmic reticulum containing high density microsomal fractions, which also contain the InsP₃Rs (Supplementary Fig. 9e). The InsP₃R–CRTC2 association seems to be critical for CRTC2 localization in the perinuclear space, because RNA interference (RNAi)-mediated knockdown of the InsP₃Rs led to redistribution of CRTC2 in the cytoplasm (Supplementary Fig. 9f). Disrupting the CRTC2–InsP₃R interaction, by deletion of the CBD in CRTC2 or by addition of an N-terminal myristoylation signal that targets CRTC2 to the plasma membrane, blocked CRTC2 dephosphorylation and CRE-reporter activation in response to glucagon (Supplementary Fig. 9g–i). Taken together, these results suggest that the association of CRTC2 with InsP₃Rs enhances its sensitivity to fasting signals.

Under feeding conditions, insulin inhibits gluconeogenesis in part by increasing CRTC2 phosphorylation. We wondered whether insulin interferes with InsP₃R effects on CRTC2 activation. Supporting this idea, AKT has been shown to block calcium mobilization by phosphorylating InsP₃Rs at Ser 2682 (in InsP₃R1)¹¹. Indeed, exposure of hepatocytes to insulin increased InsP₃R phosphorylation by immunoblot analysis with phospho-AKT substrate antiserum (Supplementary Fig. 10a); mutation of Ser 2682 (in InsP₃R1) to alanine blocked these effects. Insulin treatment also reduced glucagon-dependent increases in calcium mobilization and calcineurin activation in cells expressing wild-type InsP₃R1, but it had no effect in cells expressing AKT-defective (S2682A) InsP₃R1 (Supplementary Fig. 10b). As a result, CRTC2 dephosphorylation, CRE-Luc activity, and glucose output were elevated in hepatocytes expressing InsP₃R(S2682A) compared to wild type (Supplementary Fig. 10c).

We examined whether InsP₃R1 phosphorylation by AKT is important in regulating hepatic glucose production *in vivo*. In line with this

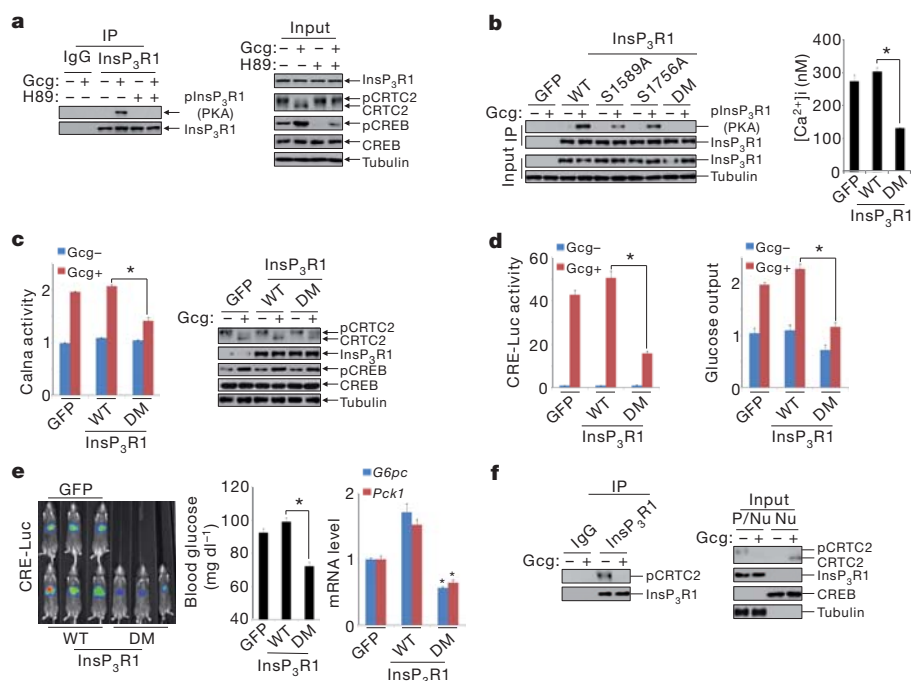


Figure 3 | Glucagon stimulates CRTC2 activity via PKA-dependent phosphorylation of InsP₃Rs. **a, b,** Immunoblots of InsP₃R1 immunoprecipitates (IP) using phospho-PKA substrate antiserum to show effect of H89 (**a**) and Ala mutations (**b**) at one or both (DM) PKA consensus sites (Ser 1589, Ser 1756) on InsP₃R1 phosphorylation in hepatocytes exposed to glucagon (Gcg). Effect of wild-type and PKA-mutant InsP₃R1 on calcium mobilization in response to Gcg (**b**) shown (**P* < 0.001; *n* = 3). **c, d,** Effect of wild-type or PKA-defective InsP₃R1 (DM) on calcineurin (Calna) activation

(**c**) and CRTC2 dephosphorylation (**c**), as well as CRE-Luc activation (**d**) and glucose output (**d**) from hepatocytes (**P* < 0.001; *n* = 3). **e,** Effect of wild-type and PKA-defective InsP₃R1 on hepatic CRE-Luc activity, fasting blood glucose, and gluconeogenic gene expression (*G6pc*, *Pck1*) (**P* < 0.01 versus wild type; *n* = 5). **f,** Co-immunoprecipitation of CRTC2 with InsP₃R1 in primary hepatocytes. Exposure to glucagon (100 nM, 15 min) indicated. Input levels of CRTC2 and InsP₃R1 in nuclear (Nu) and post-nuclear (p/Nu) supernatant fractions shown. Data are shown as mean ± s.e.m.

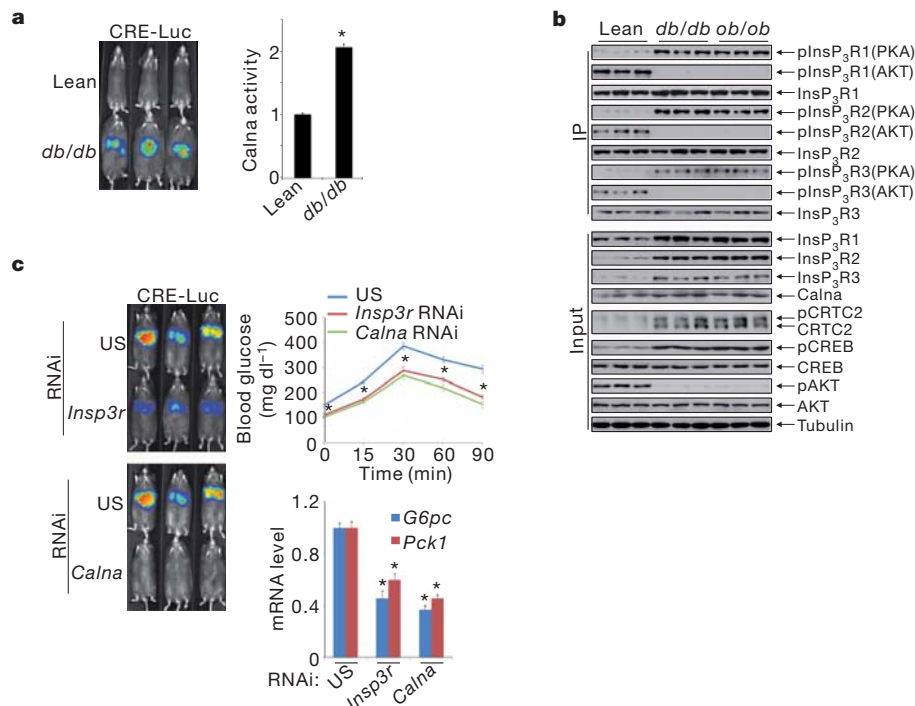


Figure 4 | InsP₃R activity is upregulated in diabetes. **a**, Hepatic CRE-Luc and calcineurin activity in lean and *db/db* mice (**P* < 0.001; *n* = 5). **b**, Immunoblots showing relative amounts and phosphorylation of InsP₃R family members in livers of ad libitum fed lean, *db/db*, or *ob/ob* mice. InsP₃R

phosphorylation at PKA or AKT sites indicated. **c**, Effect of RNAi-mediated depletion of InsP₃Rs or calcineurin A on CRE-Luc activity, gluconeogenic gene expression, and hepatic glucose production in *db/db* mice, determined by pyruvate tolerance testing (**P* < 0.01; *n* = 5). Data are shown as mean ± s.e.m.

notion, feeding increased hepatic InsP₃R1 phosphorylation at Ser 2682 (Supplementary Fig. 8b). Moreover, overexpression of AKT-defective InsP₃R1 partially suppressed feeding-induced decreases in CRE-Luc activity and gluconeogenic gene expression, leading to elevations in circulating glucose concentrations (Supplementary Fig. 10d). Taken together, these results suggest that the AKT-mediated phosphorylation of InsP₃Rs during feeding inhibits hepatic gluconeogenesis by blocking the calcineurin-dependent dephosphorylation of CRTC2.

We wondered whether hepatic InsP₃R signalling contributes to increases in gluconeogenesis in the setting of insulin resistance. Supporting this notion, hepatic calcineurin activity was enhanced in both *ob/ob* and *db/db* diabetic animals, leading to increases in CRE-Luc activity (Fig. 4a and Supplementary Fig. 11a, b). Pointing to a role for InsP₃R, hepatic amounts of PKA-phosphorylated, active InsP₃Rs were increased in these diabetic mice, whereas amounts of AKT-phosphorylated, inactive InsP₃Rs were reduced (Fig. 4b). Correspondingly, knockdown of either calcineurin or InsP₃Rs in *db/db* mice reduced CRE-Luc activity, gluconeogenic gene expression, and hepatic gluconeogenesis (Fig. 4c and Supplementary Fig. 11c).

Collectively, our results demonstrate that glucagon promotes CRTC2 dephosphorylation during fasting by triggering increases in cytoplasmic calcium that lead to calcineurin activation (Supplementary Fig. 12). The ability for glucagon to increase calcium signalling via the PKA-mediated phosphorylation of InsP₃Rs demonstrates an important regulatory node for cross-talk between cAMP and calcium signalling pathways in liver and perhaps other insulin sensitive tissues. The partial inhibition of calcium entry by the PKA inhibitor H89 also points to additional regulatory inputs^{12,13} that may function with PKA to increase InsP₃R activity in response to glucagon. CRTC2 has also been found to stimulate metabolic gene expression by upregulating the nuclear hormone receptor co-activator PGC-1 α in liver^{14,15} and muscle¹⁶. On the basis of the well-recognized role of calcium signalling in PGC-1 α -dependent transcription, InsP₃Rs may also have an important function in this setting.

METHODS SUMMARY

Adenoviruses were delivered by tail vein injection¹⁷. Hepatic CRE-Luc activity was visualized using an IVIS Imaging system. Mice were imaged 3–5 days after injection of CRE-Luc adenovirus. Pyruvate tolerance testing was performed on mice fasted overnight and injected intraperitoneally with pyruvate (2 g kg⁻¹). *Insp3r2* knockout mice have been described¹⁰. Cultured primary mouse hepatocytes were prepared as reported¹⁸. Cellular fractionation studies were conducted using primary mouse hepatocytes¹⁸. Calcium imaging experiments were performed using a CCD camera on primary hepatocytes loaded with fura-2 dye. Mass spectrometry studies were performed on CRTC2 immunoprecipitates prepared from HEK293T cells and on immunoprecipitates of phospho-PKA substrate antiserum prepared from primary hepatocytes exposed to glucagon. Anti-InsP₃R1 (A302–158A) and InsP₃R3 (A302–160A) antibodies were purchased from Bethyl Laboratories, anti-InsP₃R2 (ab77838) antiserum was from Abcam, anti-calcineurin (610260) from BD Biosciences, anti-GRP78 (ADI-SPA-826-F) from Enzo Life Sciences, anti-phospho-PKA substrate (RRXS/T, 9624), anti-phospho-AKT substrate (RRXS/T, 9614) and CRTC2 (pS171, 2892) from Cell Signaling. Phospho (Ser 275) CRTC2 antibody was used as described¹⁹. For more details, see Supplementary Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 July 2011; accepted 22 February 2012.

Published online 8 April 2012.

- Altarejos, J. Y. & Montminy, M. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. *Nature Rev. Mol. Cell Biol.* **12**, 141–151 (2011).
- Yoon, Y. S. *et al.* Suppressor of MEK null (SMEK)/protein phosphatase 4 catalytic subunit (PP4C) is a key regulator of hepatic gluconeogenesis. *Proc. Natl Acad. Sci. USA* **107**, 17704–17709 (2010).
- Screation, R. A. *et al.* The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector. *Cell* **119**, 61–74 (2004).
- Hogan, P. G., Chen, L., Nardone, J. & Rao, A. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev.* **17**, 2205–2232 (2003).
- Ferris, C. D., Haganir, R. L., Bredt, D. S., Cameron, A. M. & Snyder, S. H. Inositol trisphosphate receptor: phosphorylation by protein kinase C and calcium calmodulin-dependent protein kinases in reconstituted lipid vesicles. *Proc. Natl Acad. Sci. USA* **88**, 2232–2235 (1991).
- Volpe, P. & Alderson-Lang, B. H. Regulation of inositol 1,4,5-trisphosphate-induced Ca²⁺ release. II. Effect of cAMP-dependent protein kinase. *Am. J. Physiol.* **258**, C1086–C1091 (1990).

7. Bird, G. S., Burgess, G. M. & Putney, J. W. Jr. Sulfhydryl reagents and cAMP-dependent kinase increase the sensitivity of the inositol 1,4,5-trisphosphate receptor in hepatocytes. *J. Biol. Chem.* **268**, 17917–17923 (1993).
8. Patterson, R. L., Boehning, D. & Snyder, S. H. Inositol 1,4,5-trisphosphate receptors as signal integrators. *Annu. Rev. Biochem.* **73**, 437–465 (2004).
9. Futatsugi, A. *et al.* IP3 receptor types 2 and 3 mediate exocrine secretion underlying energy metabolism. *Science* **309**, 2232–2234 (2005).
10. Cruz, L. N. *et al.* Regulation of multidrug resistance-associated protein 2 by calcium signaling in mouse liver. *Hepatology* **52**, 327–337 (2010).
11. Szado, T. *et al.* Phosphorylation of inositol 1,4,5-trisphosphate receptors by protein kinase B/AKT inhibits Ca^{2+} release and apoptosis. *Proc. Natl Acad. Sci. USA* **105**, 2427–2432 (2008).
12. Tovey, S. C. *et al.* Regulation of inositol 1,4,5-trisphosphate receptors by cAMP independent of cAMP-dependent protein kinase. *J. Biol. Chem.* **285**, 12979–12989 (2010).
13. Wakelam, M. J., Murphy, G. J., Hruby, V. J. & Houslay, M. D. Activation of two signal-transduction systems in hepatocytes by glucagon. *Nature* **323**, 68–71 (1986).
14. Yoon, J. *et al.* Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131–138 (2001).
15. Herzig, S. *et al.* CREB regulates hepatic gluconeogenesis via the co-activator PGC-1. *Nature* **413**, 179–183 (2001).
16. Wu, Z. *et al.* Transducer of regulated CREB-binding proteins (TORCs) induce PGC-1 α transcription and mitochondrial biogenesis in muscle cells. *Proc. Natl Acad. Sci. USA* **103**, 14379–14384 (2006).
17. Dentin, R. *et al.* Insulin modulates gluconeogenesis by inhibition of the coactivator TORC2. *Nature* **449**, 366–369 (2007).
18. Wang, Y., Vera, L., Fischer, W. H. & Montminy, M. The CREB coactivator CRTC2 links hepatic ER stress and fasting gluconeogenesis. *Nature* **460**, 534–537 (2009).
19. Jansson, D. *et al.* Glucose controls CREB activity in islet cells via regulated phosphorylation of TORC2. *Proc. Natl Acad. Sci. USA* **105**, 10161–10166 (2008).
20. Liu, Y. *et al.* A fasting inducible switch modulates gluconeogenesis via activator/coactivator exchange. *Nature* **456**, 269–273 (2008).
21. Wang, B. *et al.* A hormone-dependent module regulating energy balance. *Cell* **145**, 596–606 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by National Institutes of Health grants R01-DK049777, R01-DK083834 and R01-DK091618 (M.M.), HL087123 (I.T.), the Kieckhefer Foundation, The Clayton Foundation for Medical Research, and the Leona M. and Harry B. Helmsley Charitable Trust.

Author Contributions Y.W., I.T. and M.M. designed and interpreted the experiments. Y.W., G.L., J.C.P., R.S. and J.G. carried out the experimental work. Y.W., K.O. and J.C. characterized glucose metabolism in *Insp3r2* knockout mice. W.H.F. performed proteomic studies, and Y.W. and M.M. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.M. (montminy@salk.edu).

METHODS

Mouse strains and adenovirus. Adenoviruses (1×10^8 plaque forming units (p.f.u.) GFP, calcineurin, *InsP₃R1*, *InsP₃R1* DM (S1589A/S1756A), unspecific RNAi, calcineurin RNAi, *Insp3r1* RNAi, *Insp3r2* RNAi, *Insp3r3* RNAi, *Crtc2* RNAi, 1×10^9 p.f.u. CRE-Luc reporter, 5×10^7 p.f.u. RSV β -galactosidase) were delivered to 8–10-week-old male C57BL/6J, B6.V-*lep* < ob>/J, B6.Cg-*m*+/+*Lepr* < db>/J mice by tail vein injection¹⁷. *Insp3r2* knockout mice were described previously¹⁰. All mice were adapted to their environment for 1 week before study and were housed in colony cages with a 12 h light/dark cycle in a temperature-controlled environment. For *in vivo* imaging experiments, mice were imaged on day 3–5 after adenovirus delivery. Wild-type CRTC2, CRTC2(S171A), GFP, unspecific RNAi, *Crtc2* RNAi, CRE-Luc and RSV β -gal adenoviruses have been described previously^{17,20}. The adenoviruses containing rat *InsP₃R1*, *InsP₃R1* DM and *InsP₃R1*(S2682A) were generated from the *InsP₃R1* plasmid, provided by I. Bezprozvanny (UT Southwestern Medical Center at Dallas). Calcineurin adenovirus was constructed using a mouse calcineurin plasmid (Addgene). CRTC2 Δ CBD (51–692 amino acids), S275A and S171A/S275A adenoviruses were made from mouse CRTC2. Myristoylated CRTC2 (Myr-CRTC2) adenovirus was generated with mouse CRTC2 fused to an N-terminal myristoylation tag (MGSSKSKPKDPSQR) from Src. Calcineurin RNAi, *Insp3r1* RNAi, *Insp3r2* RNAi, *Insp3r3* RNAi adenoviruses were constructed using the sequence 5'-GGGTACCGCATGTACAGGAAAA-3', 5'-GGGTACTGGAATAGCCTCTTCC-3', 5'-GGGTAACAAGCACCACCATCCCC-3' and 5'-GGGCAAGCTGCA GGTGTTCTCTG-3', respectively. All expressed constructs used in this study were confirmed by sequencing.

In vivo analysis. For *in vivo* imaging, mice were imaged as described^{17,20} under ad libitum feeding conditions or after fasting for 6 h. For pyruvate challenge experiments, mice were fasted overnight and injected intraperitoneally with pyruvate (2 g kg^{-1}). Blood glucose values were determined using a LifeScan automatic glucometer. For immunoblot, mouse tissues were sonicated, centrifuged and supernatants were reserved for protein determinations, and SDS-PAGE analysis.

In vitro analysis. HEK293T (ATCC) cells were cultured in DMEM containing 10% FBS (HyClone), 100 mg ml^{-1} penicillin-streptomycin. Mouse primary hepatocytes were isolated and cultured as previously described¹⁸. Cellular fractionation studies were conducted as previously reported¹⁸. For reporter studies, Ad-CRE-Luc-infected hepatocytes ($1 \text{ p.f.u. per cell}$) were exposed to glucagon (100 nM) for 2 to $\sim 4 \text{ h}$. For CsA ($10 \text{ }\mu\text{M}$), okadaic acid (100 nM), cell permeable calcineurin autoinhibitory peptide ($10 \text{ }\mu\text{M}$), CN585 ($100 \text{ }\mu\text{M}$), calyculin A (10 nM), xestospongin C ($2 \text{ }\mu\text{M}$), H89 ($30 \text{ }\mu\text{M}$) or BAPTA ($50 \text{ }\mu\text{M}$) inhibition, hepatocytes were pre-treated with the inhibitors for 1 h. Luciferase activities were normalized to β -galactosidase activity from adenoviral-encoded RSV β -galactosidase.

Calcineurin activity (test kit from Enzo Life Sciences) and cellular cAMP levels (test kit from Cayman Chemical Company) were measured according to the manufacturer's instructions.

Calcium imaging. Mouse primary hepatocytes were plated on glass coverslips and loaded with $5 \text{ }\mu\text{M}$ Fura-2 acetoxymethyl ester (Molecular Probes) in the presence of 0.025% (w/v) pluronic F127 (Sigma-Aldrich) in Media 199 (Mediatech) for 30 min. Coverslips were mounted on a laminar flow perfusion chamber (Warner Instruments) and perfused with Media 199 or a solution of 100 nM glucagon in Media 199. Images of Fura-2 loaded cells were collected with a cooled CCD camera while the excitation wavelength was alternated between 340 nm and 380 nm. The ratio of fluorescence intensity at the two excitation wavelengths was calculated after subtracting background fluorescence. $[\text{Ca}^{2+}]_i$ (cytosolic free calcium concentration) was calculated using a Fura-2 calcium imaging calibration kit (Invitrogen). Images were collected and analysed using the MetaFluor software package (Universal Imaging Corp.). Graphs represent average responses from groups of 30–40 individual cells from representative single experiments. Bar graphs represent average responses (fold over average baseline) from 150–200 cells per condition. All experiments were repeated at least three times with similar results.

Immunoblot, immunoprecipitation and immunostaining. Immunoblot, immunoprecipitation and immunostaining assays were performed as described¹⁸. CRTC2, pCREB (Ser 133), CREB, pAKT (Thr 308), AKT, tubulin, HA and Flag antibodies were previously described¹⁸. The antibodies anti-*InsP₃R1* (A302–158A) and *InsP₃R3* (A302–160A) were purchased from Bethyl Laboratories, anti-*InsP₃R2* (ab77838) from Abcam, anti-calcineurin (610260) from BD Biosciences, anti-GRP78 (ADI-SPA-826-F) from Enzo Life Sciences, anti-phospho-PKA substrate (RRXS/T, 9624), anti-phospho-AKT substrate (RRXS/T, 9614) and CRTC2 (pS171, 2892) from Cell Signaling. CRTC2 (pS275) antibody was used as described¹⁹.

Quantitative PCR. Total cellular RNAs from whole liver or from primary hepatocytes were extracted using the RNeasy kit (Qiagen) and used to generate cDNA with SuperScript II enzyme (Invitrogen). cDNA were analysed by quantitative PCR as described¹⁸.

Mass spectrometry. Immunoprecipitates of endogenous CRTC2 from HEK293T cells and of phospho-PKA substrate antiserum from glucagon-stimulated hepatocytes were prepared for mass spectrometric studies as previously reported²¹, and analysed by electrospray ionization tandem mass spectrometry on a Thermo LTQ Orbitrap instrument.

Statistical analyses. All studies were performed on at least three independent occasions. Results are reported as mean \pm s.e.m. The comparison of different groups was carried out using two-tailed unpaired Student's *t*-test. Differences were considered statistically significant at $P < 0.05$.

Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations

Brian J. O’Roak¹, Laura Vives¹, Santhosh Girirajan¹, Emre Karakoc¹, Niklas Krumm¹, Bradley P. Coe¹, Roie Levy¹, Arthur Ko¹, Choli Lee¹, Joshua D. Smith¹, Emily H. Turner¹, Ian B. Stanaway¹, Benjamin Vernot¹, Maika Malig¹, Carl Baker¹, Beau Reilly², Joshua M. Akey¹, Elhanan Borenstein^{1,3,4}, Mark J. Rieder¹, Deborah A. Nickerson¹, Raphael Bernier², Jay Shendure¹ & Evan E. Eichler^{1,5}

It is well established that autism spectrum disorders (ASD) have a strong genetic component; however, for at least 70% of cases, the underlying genetic cause is unknown¹. Under the hypothesis that *de novo* mutations underlie a substantial fraction of the risk for developing ASD in families with no previous history of ASD or related phenotypes—so-called sporadic or simplex families^{2,3}—we sequenced all coding regions of the genome (the exome) for parent–child trios exhibiting sporadic ASD, including 189 new trios and 20 that were previously reported⁴. Additionally, we also sequenced the exomes of 50 unaffected siblings corresponding to these new ($n = 31$) and previously reported trios ($n = 19$)⁴, for a total of 677 individual exomes from 209 families. Here we show that *de novo* point mutations are overwhelmingly paternal in origin (4:1 bias) and positively correlated with paternal age, consistent with the modest increased risk for children of older fathers to develop ASD⁵. Moreover, 39% (49 of 126) of the most severe or disruptive *de novo* mutations map to a highly interconnected β -catenin/chromatin remodelling protein network ranked significantly for autism candidate genes. In proband exomes, recurrent protein-altering mutations were observed in two genes: *CHD8* and *NTNG1*. Mutation screening of six candidate genes in 1,703 ASD probands identified additional *de novo*, protein-altering mutations in *GRIN2B*, *LAMC3* and *SCN1A*. Combined with copy number variant (CNV) data, these results indicate extreme locus heterogeneity but also provide a target for future discovery, diagnostics and therapeutics.

We selected 189 autism trios from the Simons Simplex Collection (SSC)⁶, which included males significantly impaired with autism and intellectual disability ($n = 47$), a female sample set ($n = 56$) of which 26 were cognitively impaired, and samples chosen at random from the remaining males in the collection ($n = 86$) (Supplementary Table 1 and Supplementary Fig. 1). In general, we excluded samples known to carry large *de novo* CNVs². Exome sequencing was performed as described previously⁴, but with an expanded target definition (see Methods). We achieved sufficient coverage for both parents and child to call genotypes for, on average, 29.5 megabases (Mb) of haploid exome coding sequence (Supplementary Table 1). In addition, we performed copy number analysis on 122 of these families, using a combination of the exome data, array comparative genomic hybridization (CGH), and genotyping arrays, thereby providing a more comprehensive view of rare variation.

In the 189 new probands, we validated 248 *de novo* events, 225 single nucleotide variants (SNVs), 17 small insertions/deletions (indels), and six CNVs (Supplementary Table 2). These included 181 non-synonymous changes, of which 120 were classified as severe based on sequence conservation and/or biochemical properties (Methods and Supplementary Table 3). The observed point mutation rate in coding sequence was ~ 1.3 events per trio or 2.17×10^{-8} per base

per generation, in close agreement with our previous observations⁴, yet in general, higher than previous studies, indicating increased sensitivity (Supplementary Table 2 and Supplementary Table 4)⁷. We also observed complex classes of *de novo* mutation including: five cases of multiple mutations in close proximity; two events consistent with paternal germline mosaicism (that is, where both siblings contained a *de novo* event observed in neither parent); and nine events showing a weak minor allele profile consistent with somatic mosaicism (Supplementary Table 3 and Supplementary Figs 2 and 3).

Of the severe *de novo* events, 28% (33 of 120) are predicted to truncate the protein. The distribution of synonymous, missense and nonsense changes corresponds well with a random mutation model⁷ (Supplementary Fig. 4 and Supplementary Table 2). However, the difference in nonsense rates between *de novo* and rare singleton events (not present in 1,779 other exomes) is striking (4:1) and suggests strong selection against new nonsense events (Fisher’s exact test, $P < 0.0001$). In contrast with a recent report⁸, we find no significant difference in mutation rate between affected and unaffected individuals; however, we do observe a trend towards increased non-synonymous rates in probands, consistent with the findings of ref. 9 (Supplementary Tables 1 and 2).

Given the association of ASD with increased paternal age⁵ and our previous observations⁴, we used molecular cloning, read-pair information, and obligate carrier status to identify informative markers linked to 51 *de novo* events and observed a marked paternal bias (41:10; binomial $P < 1.4 \times 10^{-5}$; Fig. 1a and Supplementary Tables 3 and 5). This provides strong direct evidence that the germline mutation rate in protein-coding regions is, on average, substantially higher in males. A similar finding was recently reported for *de novo* CNVs¹⁰. In addition, we observe that the number of *de novo* events is positively correlated with increasing paternal age (Spearman’s rank correlation = 0.19; $P < 0.008$; Fig. 1b). Together, these observations are consistent with the hypothesis that the modest increased risk for children of older fathers to develop ASD⁵ is the result of an increased mutation rate.

Using sequence read-depth methods in 122 of the 189 families, we scanned ASD probands for either *de novo* CNVs or rare ($< 1\%$ of controls), inherited CNVs. Individual events were validated by either array CGH or genotyping array (see Methods). We identified 76 events in 53 individuals, including six *de novo* (median size 467 kilobases (kb)) and 70 inherited (median size 155 kb) CNVs (Supplementary Table 6). These include disruptions of *EHMT1* (Kleefstra’s syndrome, Online Mendelian Inheritance in Man (OMIM) accession 610253), *CNTNAP4* (reported in children with developmental delay and autism¹¹) and the 16p11.2 duplication (OMIM 611913) associated with developmental delay, bipolar disorder and schizophrenia.

We performed a multivariate analysis on non-verbal IQ (NVIQ), verbal IQ (VIQ) and the load of ‘extreme’ *de novo* mutations—where extreme is defined as point mutations that truncate proteins, intersect

¹Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA. ²Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington 98195, USA. ³Department of Computer Science and Engineering, University of Washington, Seattle, Washington 98195, USA. ⁴Santa Fe Institute, Santa Fe, New Mexico 87501, USA. ⁵Howard Hughes Medical Institute, Seattle, Washington 98195, USA.

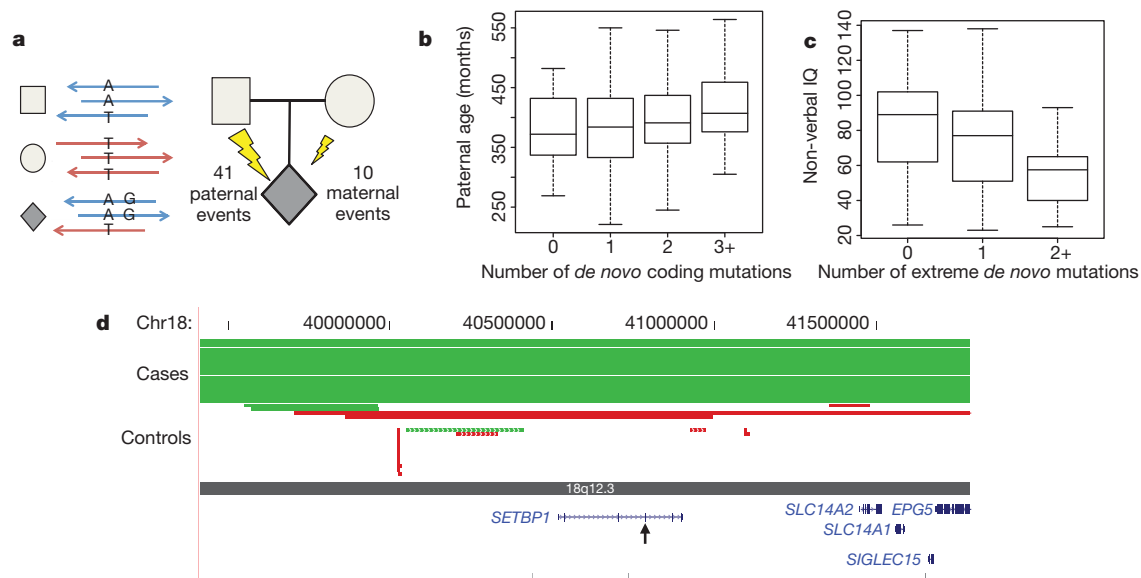


Figure 1 | De novo mutation events in autism spectrum disorder.

a, Haplotype phasing using informative markers shows a strong parent-of-origin bias with 41 of 51 *de novo* events occurring on the paternally inherited haplotype. Arrows represent sequence reads from paternal (blue) or maternal (red) haplotypes. **b**, **c**, Box and whisker plots for 189 SSC probands. **b**, The paternal estimated age at conception versus the number of observed *de novo* point mutations (0, $n = 53$; 1, $n = 65$; 2, $n = 44$; 3+, $n = 27$). **c**, Decreased non-verbal IQ is significantly associated with an increasing number of extreme

mutation events (0, $n = 138$; 1, $n = 41$; 2+, $n = 10$), both with and without CNVs (Supplementary Discussion). **d**, Browser images showing CNVs identified in the del(18)(q12.2q21.1) syndrome region. The truncating point mutation in *SETBP1* occurs within the critical region, identifying the likely causative locus. Each red (deletion) and green (duplication) line represents an identified CNV in cases (solid lines) versus controls (dashed lines), with arrowheads showing point mutation.

Mendelian or ASD loci ($n = 57$), or *de novo* CNVs that intersect genes ($n = 5$) (Fig. 1c and Supplementary Discussion). NVIQ, but not VIQ, decreased significantly ($P < 0.01$) with increased number of events. Covariant analysis of the samples with CNV data showed that this finding was strengthened, but not exclusively driven, by the presence of either *de novo* or rare CNVs (Supplementary Fig. 5).

Among the *de novo* events, we identified 62 top ASD risk contributing mutations based on the deleteriousness of the mutations, functional evidence, or previous studies (Table 1). Probands with these mutations spanned the range of IQ scores, with only a modest non-significant trend towards individual's co-morbid with intellectual disability (Supplementary Figs 1 and 6). We observed recurrent, protein-disruptive mutations in two genes: *NTNG1* (netrin G1) and *CHD8* (chromodomain helicase DNA binding protein 8). Given their locus-specific mutation rates, the probability of identifying two independent mutations in our sample set is low (uncorrected, *NTNG1*: $P < 1.2 \times 10^{-6}$; *CHD8*: $P < 6.9 \times 10^{-5}$) (Supplementary Fig. 7, Supplementary Table 8 and Methods). *NTNG1* is a strong biological candidate given its role in laminar organization of dendrites and axonal guidance¹² and was also reported as being disrupted by a *de novo* translocation in a child with Rett's syndrome, without *MECP2* mutation¹³. Both *de novo* mutations identified here are missense (p.Tyr23Cys and p.Thr135Ile) at highly conserved positions predicted to disrupt protein function, although there is evidence of mosaicism for the former mutation (Supplementary Table 3).

CHD8 has not previously been associated with ASD and codes for an ATP-dependent chromatin-remodelling factor that has a significant role in the regulation of both β -catenin and p53 signalling^{14,15}. We also identified *de novo* missense variants in *CHD3* as well as *CHD7* (CHARGE syndrome, OMIM 214800), a known binding partner of *CHD8* (ref. 16). ASD has been found in as many as two-thirds of children with CHARGE, indicating that *CHD7* may contribute to an ASD syndromic subtype¹⁷.

We identified 30 protein-altering *de novo* events intersecting with Mendelian disease loci (Supplementary Table 3) as well as inherited hemizygous mutations of clinical significance (Supplementary Table 9).

The *de novo* mutations included truncating events in syndromic intellectual disability genes (*MBD5* (mental retardation, autosomal dominant 1, OMIM 156200), *RPS6KA3* (Coffin–Lowry syndrome, OMIM 303600) and *DYRK1A* (the Down's syndrome candidate gene, OMIM 600855)), and missense variants in loci associated with syndromic ASD, including *CHD7*, *PTEN* (macrocephaly/autism syndrome, OMIM 605309) and *TSC2* (tuberous sclerosis complex, OMIM 613254). Notably, *DYRK1A* is a highly conserved gene mapping to the Down's syndrome critical region (Supplementary Fig. 8). The proband here (13890) is severely cognitively impaired and microcephalic, consistent with previous studies of *DYRK1A* haploinsufficiency in both patients and mouse models¹⁸.

Twenty-one of the non-synonymous *de novo* mutations map to CNV regions recurrently identified in children with developmental delay and ASD (Supplementary Table 10), such as *MBD5* (2q23.1 deletion syndrome), *SYNRG* (17q12 deletion syndrome) and *POLRMT* (19p13.3 deletion)¹⁹. There is also considerable overlap with genes disrupted by single *de novo* CNVs in children with ASD (for example, *NLGN1* and *ARID1B*; Supplementary Table 11). Given the prior probability that these loci underlie genomic disorders, the disruptive *de novo* SNVs and small indels may be pinpointing the possible major effect locus for ASD-related features. For example, we identified a complex *de novo* mutation resulting in truncation of *SETBP1* (SET binding protein 1), one of five genes in the critical region for del(18)(q12.2q21.1) syndrome (Fig. 1d), which is characterized by hypotonia, expressive language delay, short stature and behavioural problems²⁰. Recurrent *de novo* missense mutations at *SETBP1* were recently reported to be causative for a distinct phenotype, Schinzel–Giedion syndrome, probably through a gain-of-function mechanism²¹, indicating diverse phenotypic outcomes at this locus depending on mutation mechanism.

Several of the mutated genes encode proteins that directly interact, suggesting a common biological pathway. From our full list of genes carrying truncating or severe missense mutations (126 events from all 209 families), we generated a protein–protein interaction (PPI) network based on a database of physical interactions (Supplementary Table 12)²². We found 39% (49 of 126) of the genes mapped to a highly

Table 1 | Top *de novo* ASD risk contributing mutations

| Proband | NVIQ | Candidate gene | Amino acid change |
|----------|------|-----------------|--------------------------|
| 12225.p1 | 89 | <i>ABCA2</i> | p.Val1845Met |
| 11653.p1 | 44 | <i>ADCY5</i> | p.Arg603Cys |
| 12130.p1 | 55 | <i>ADNP</i> | Frameshift indel |
| 11224.p1 | 112 | <i>AP3B2</i> | p.Arg435His |
| 13447.p1 | 51 | <i>ARID1B</i> | Frameshift indel |
| 13415.p1 | 48 | <i>BRSK2</i> | 3n indel |
| 14292.p1 | 49 | <i>BRWD1</i> | Frameshift indel |
| 11872.p1 | 65 | <i>CACNA1D</i> | p.Ala769Gly |
| 11773.p1 | 50 | <i>CACNA1E</i> | p.Gly1209Ser |
| 13606.p1 | 60 | <i>CDC42BPB</i> | p.Arg764TERM |
| 12086.p1 | 108 | <i>CDH5</i> | p.Arg545Trp |
| 12630.p1 | 115 | <i>CHD3</i> | p.Arg1818Trp |
| 13733.p1 | 68 | <i>CHD7</i> | p.Gly996Ser |
| 13844.p1 | 34 | <i>CHD8</i> | p.Gln959TERM |
| 12752.p1 | 93 | <i>CHD8</i> | Frameshift indel |
| 13415.p1 | 48 | <i>CNOT4</i> | p.Asp48Asn |
| 12703.p1 | 58 | <i>CTNNB1</i> | p.Thr551Met |
| 11452.p1 | 80 | <i>CUL3</i> | p.Glu246TERM |
| 11571.p1 | 94 | <i>CUL5</i> | p.Val355Ile |
| 13890.p1 | 42 | <i>DYRK1A</i> | Splice site |
| 12741.p1 | 87 | <i>EHD2</i> | p.Arg167Cys |
| 11629.p1 | 67 | <i>FBXO10</i> | p.Glu54Lys |
| 13629.p1 | 63 | <i>GPS1</i> | p.Arg492Gln |
| 13757.p1 | 91 | <i>GRINL1A</i> | 3n indel |
| 11184.p1 | 94 | <i>HDGFRP2</i> | p.Glu83Lys |
| 11610.p1 | 138 | <i>HDLBP</i> | p.Ala639Ser |
| 11872.p1 | 65 | <i>KATNAL2</i> | Splice site |
| 12346.p1 | 77 | <i>MBD5</i> | Frameshift indel |
| 11947.p1 | 33 | <i>MDM2</i> | p.Glu433Lys/p.Trp160TERM |
| 11148.p1 | 82 | <i>MLL3</i> | p.Tyr4691TERM |
| 12157.p1 | 91 | <i>NLGN1</i> | p.His795Tyr |
| 11193.p1 | 138 | <i>NOTCH3</i> | p.Gly1134Arg |
| 11172.p1 | 60 | <i>NR4A2</i> | p.Tyr275His |
| 11660.p1 | 60 | <i>NTNG1</i> | p.Thr135Ile |
| 12532.p1 | 110 | <i>NTNG1</i> | p.Tyr23Cys |
| 11093.p1 | 91 | <i>OPRL1</i> | p.Arg157Cys |
| 13793.p1 | 56 | <i>PCDH4</i> | p.Asp555His |
| 11707.p1 | 23 | <i>PDCD1</i> | Frameshift indel |
| 12304.p1 | 83 | <i>PSEN1</i> | p.Thr421Ile |
| 11390.p1 | 77 | <i>PTEN</i> | p.Thr167Asn |
| 13629.p1 | 63 | <i>PTPRK</i> | p.Arg784His |
| 13333.p1 | 69 | <i>RGMA</i> | p.Val379Ile |
| 13222.p1 | 86 | <i>RPS6KA3</i> | p.Ser369TERM |
| 11257.p1 | 128 | <i>RUVBL1</i> | p.Leu365Gln |
| 11843.p1 | 113 | <i>SESN2</i> | p.Ala46Thr |
| 12933.p1 | 41 | <i>SETBP1</i> | Frameshift indel |
| 12565.p1 | 79 | <i>SETD2</i> | Frameshift indel |
| 12335.p1 | 47 | <i>TBL1XR1</i> | p.Leu282Pro |
| 11480.p1 | 41 | <i>TBR1</i> | Frameshift indel |
| 11569.p1 | 67 | <i>TNKS</i> | p.Arg568Thr |
| 12621.p1 | 120 | <i>TSC2</i> | p.Arg1580Trp |
| 11291.p1 | 83 | <i>TSPAN17</i> | p.Ser75TERM |
| 11006.p1 | 125 | <i>UBE3C</i> | p.Ser845Phe |
| 12161.p1 | 95 | <i>UBR3</i> | Frameshift indel |
| 12521.p1 | 78 | <i>USP15</i> | Frameshift indel |
| 11526.p1 | 92 | <i>ZBTB41</i> | p.Tyr886His |
| 13335.p1 | 25 | <i>ZNF420</i> | p.Leu76Pro |

| Proband | NVIQ | CNV Candidate gene | Type |
|----------|------|-----------------------|-------------|
| 11928.p1 | 66 | <i>CHRNA7</i> | Duplication |
| 13815.p1 | 56 | <i>CNTNAP4</i> | Deletion |
| 13726.p1 | 59 | <i>CTNND1</i> | Deletion |
| 12581.p1 | 34 | <i>EHMT1</i> | Deletion |
| 13335.p1 | 25 | <i>TBX6</i> | Duplication |

Top candidate mutations based on severity and/or supporting evidence from the literature.

interconnected network wherein 92% of gene pairs in the connected component are linked by paths of three or fewer edges (Fig. 2a). We tested this degree of interconnectivity by simulation ($n = 10,000$ replicates; Methods and Supplementary Fig. 9) and found that our experimental network had significantly more edges ($P < 0.0001$) and a greater clustering coefficient ($P < 0.0001$) than expected by chance.

To investigate the relevance of this network to autism further, we applied degree-aware disease gene prioritization (DADA)²³, based on the same PPI database to rank all genes based on their relatedness to a

set of 103 previously identified ASD genes¹⁷. We found that the genes with severe mutations ranked significantly higher than all other genes (Mann–Whitney U -test, $P < 4.0 \times 10^{-4}$), suggesting enrichment of ASD candidates. Furthermore, the 49 members of the connected component overwhelmingly drove this difference (Mann–Whitney U -test, $P < 1.6 \times 10^{-8}$), as the unconnected members were not significant on their own (Mann–Whitney U -test, $P < 0.28$), increasing our confidence that these connected gene products are probably related to ASD (Supplementary Fig. 10). Consistent with this finding, the rankings of unaffected sibling events are highly similar to the unconnected component, strengthening our confidence in the enrichment of the connected component of proband events for ASD-relevant genes.

Members of this network have known functions in β -catenin and p53 signalling, chromatin remodelling, ubiquitination and neuronal development (Fig. 2a). A fundamental developmental regulator observed in the network is *CTNNB1* (catenin (cadherin-associated protein), $\beta 1$, 88 kDa), also known as β -catenin. Interestingly, a parallel analysis using ingenuity pathway analysis (IPA) shows an enrichment of upstream interacting genes of the β -catenin pathway (8 of 358, $P = 0.0030$; see Methods, Supplementary Table 13 and Supplementary Fig. 11). A role for Wnt/ β -catenin signalling in ASD was previously proposed²⁴, largely on the basis of the association of common variants in *EN2* and *WNT2*, and the high rate of children with macrocephaly. It is striking that both individuals with *CHD8* mutations in this study have multiple *de novo* disruptive missense mutations in this pathway or closely related pathways (Fig. 2b, c and Supplementary Fig. 12) and both have macrocephaly.

In addition, the pathway analysis shows several other disrupted genes not identified in the PPI that are involved in common pathways, which in some cases are linked to β -catenin (Supplementary Discussion and Supplementary Fig. 11). *TBR1*, for example, is a transcription factor that has a critical role in the development of the cerebral cortex²⁵. *TBR1* binds with *CASK* and regulates several candidate genes for ASD and intellectual disability including *GRIN2B*, *AUTS2* and *RELN*—genes of recurrent ASD mutation, some of which are described here and in other studies^{4,9,11,17}.

Our exome analysis of *de novo* coding mutations in 209 autism trios identified only two recurrently altered genes, consistent with extreme locus heterogeneity underlying ASD. This extreme heterogeneity necessitates the analysis of very large cohorts for validation. We implemented a cost-effective approach based on molecular inversion probe (MIP) technology²⁶ for the targeted resequencing of six candidate genes in ~2,500 individuals, including 1,703 simplex ASD probands and 744 controls. Four of these candidates (*FOXP1*, *GRIN2B*, *LAMC3* and *SCN1A*) were identified previously⁴, whereas two (*FOXP2*, OMIM 602081 and *GRIN2A*, OMIM 613971) are related genes implicated in other neurodevelopmental phenotypes. We identified all previously observed *de novo* events (that is, in the same individuals), as well as additional *de novo* events in *GRIN2B* (two protein-truncating events), *SCN1A* (a missense) and *LAMC3* (a missense) (Supplementary Table 8). The observed number of *de novo* events was compared with expectations based on the mutation rates estimated for each gene (Methods and Supplementary Table 8), with *GRIN2B* showing the highest significance (uncorrected P value < 0.0002). Notably, the three *de novo* events observed in *GRIN2B* are all predicted to be protein truncating, whereas no events truncating *GRIN2B* were found in more than 3,000 controls (Methods).

Our analysis predicts extreme locus heterogeneity underlying the genetic aetiology of autism. Under a strict sporadic disorder-*de novo* mutation model, if 20–30% of our *de novo* point mutations are considered to be pathogenic, we can estimate between 384 and 821 loci (Methods and Supplementary Fig. 13). We reach a similar estimate if we consider recurrences from ref. 9. It is clear from phenotype and genotype data that there are many ‘autisms’ represented under the current umbrella of ASD and other genetic models are more likely in different contexts (for example, families with multiple affected

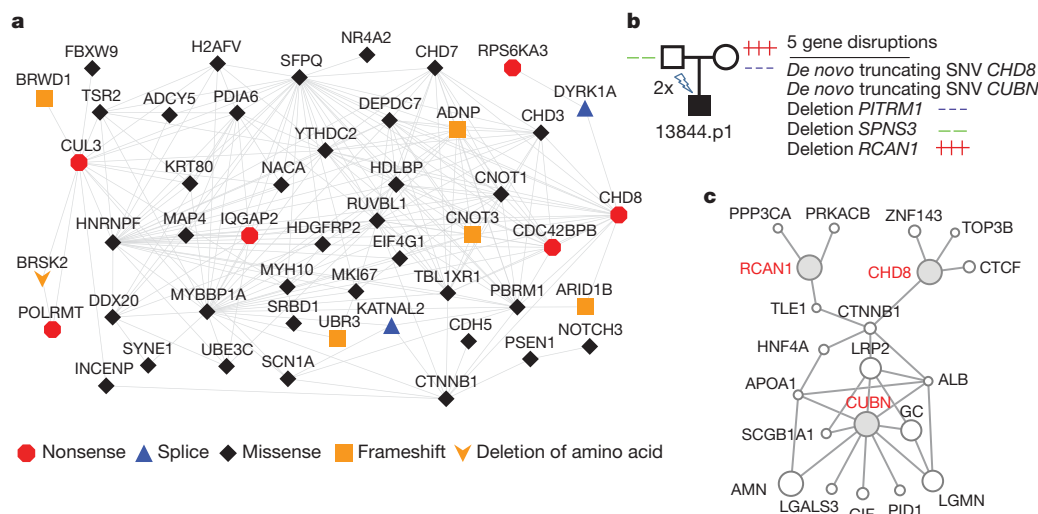


Figure 2 | Mutations identified in protein–protein interaction (PPI) networks. **a**, The 49-gene connected component of the PPI network formed from 126 genes with severe *de novo* mutations among the 209 probands. **b**, Proband 13844 inherits three rare gene-disruptive CNVs and carries two *de*

individuals). There is marked convergence on genes previously implicated in intellectual disability and developmental delay. As has been noted for CNVs, this indicates that nosological divisions may not readily translate into differences at the molecular level. We believe that there is value in comparing mutation patterns in children with developmental delay (without features of autism) to those in children with ASD.

Although there is no one major genetic lesion responsible for ASD, it is still largely unknown whether there are subsets of individuals with a common or strongly related molecular aetiology and how large these subsets are likely to be. Using gene expression, protein–protein interactions, and CNV pathway analysis, recent reports have highlighted the role of synapse formation and maintenance^{27–29}. We find it intriguing that 49 proteins found to be mutated here have critical roles in fundamental developmental pathways, including β -catenin and p53 signalling, and that patients have been identified with multiple disruptive *de novo* mutations in interconnected pathways. The latter observations are consistent with an oligogenic model of autism where both *de novo* and extremely rare inherited SNV and CNV mutations contribute in conjunction to the overall genetic risk. Recent work has supported a role for these interconnected pathways in neuronal stem-cell fate-determination, differentiation and synaptic formation in humans and animal models^{24,30,31}. Given that fundamental developmental processes have previously been found to underlie syndromic forms of autism, a wider role of these pathways in idiopathic ASD would not be entirely surprising and would help explain the extreme genetic heterogeneity observed in this study.

METHODS SUMMARY

Exome capture, alignments and base-calling. Genomic DNA was derived directly from whole blood. Exomes were considered to be completed when ~90% of the capture target exceeded 8-fold coverage and ~80% exceeded 20-fold coverage. Exomes for the 189 trios (and 31 unaffected siblings) were captured with NimbleGen EZ Exome V2.0. Reads were mapped as in ref. 4 to a custom reference genome assembly (GRC build37). Genotypes were generated with GATK unified genotyper and parallel SAMtools pipeline⁴. Exomes for the unaffected siblings matching the pilot trios were captured and analysed as in ref. 4. Predicted *de novo* events were called as in ref. 4 and confirmed by capillary sequencing in all family members (for 176 of the 189 trios, this also included one unaffected sibling). Mutations were considered severe if they were truncating, missense with Grantham score ≥ 50 and GERP score ≥ 3 or only Grantham score ≥ 85 , or deleted a highly conserved amino acid.

Exome read-depth CNV analysis. Reads were mapped using mrsFAST and normalized reads per kilobase of exon per million mapped reads (RPKM) values

novo truncating mutations. **c**, GeneMANIA²² view of three of the affected genes (**b**) (red labels) which encode proteins that are part of a β -catenin-linked network. This proband is macrocephalic, impaired cognitively, and has deficits in social behaviour and language development (Supplementary Discussion).

calculated by exon. Population normalization was performed using a set of 366 non-ASD exomes. Calls were made if three or more exons passed a threshold value and cross-validated calls using two orthogonal platforms, custom array CGH and Illumina 1M array data². CNVs were filtered to identify *de novo* and rare inherited events by comparison with 2,090 controls and 1,651 parent profiles.

Network reconstruction and null model estimation. PPI networks were generated using physical interaction data from GeneMANIA²². Null models were estimated using gene-specific mutation rate estimates based on human–chimp divergence. To rank candidate genes we obtained the seed ASD list from ref. 17 and severe disruptive *de novo* events from all families ($n = 209$). Given the PPI network and seed gene product list, we used DADA²³ for ranking each gene.

Human subjects. All samples and phenotypic data were collected under the direction of the Simons Simplex Collection by its 12 research clinic sites (<http://sfari.org/sfari-initiatives/simons-simplex-collection>). Parents consented and children assented as required by each local institutional review board. Participants were de-identified before distribution. Research was approved by the University of Washington Human Subject Division under non-identifiable biological specimens/data.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 September 2011; accepted 23 February 2012.

Published online 4 April 2012.

1. Schaaf, C. P. & Zoghbi, H. Y. Solving the autism puzzle a few pieces at a time. *Neuron* **70**, 806–808 (2011).
2. Sanders, S. J. *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70**, 863–885 (2011).
3. Levy, D. *et al.* Rare *de novo* and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **70**, 886–897 (2011).
4. O’Roak, B. J. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nature Genet.* **43**, 585–589 (2011).
5. Hultman, C. M., Sandin, S., Levine, S. Z., Lichtenstein, P. & Reichenberg, A. Advancing paternal age and risk of autism: new evidence from a population-based study and a meta-analysis of epidemiological studies. *Mol. Psychiatry* **16**, 1203–1212 (2010).
6. Fischbach, G. D. & Lord, C. The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron* **68**, 192–195 (2010).
7. Lynch, M. Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl Acad. Sci. USA* **107**, 961–968 (2010).
8. Xu, B. *et al.* Exome sequencing supports a *de novo* mutational paradigm for schizophrenia. *Nature Genet.* **43**, 864–868 (2011).
9. Sanders, S. J. *et al.* *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* <http://dx.doi.org/10.1038/nature10945> (this issue).
10. Hehir-Kwa, J. Y. *et al.* *De novo* copy number variants associated with intellectual disability have a paternal origin and age bias. *J. Med. Genet.* **48**, 776–778 (2011).
11. O’Roak, B. J. & State, M. W. Autism genetics: strategies, challenges, and opportunities. *Autism Res.* **1**, 4–17 (2008).

12. Nishimura-Akiyoshi, S., Niimi, K., Nakashiba, T. & Itohara, S. Axonal netrin-Gs transneuronally determine lamina-specific subdendritic segments. *Proc. Natl Acad. Sci. USA* **104**, 14801–14806 (2007).
13. Borg, I. *et al.* Disruption of Netrin G1 by a balanced chromosome translocation in a girl with Rett syndrome. *Eur. J. Hum. Genet.* **13**, 921–927 (2005).
14. Nishiyama, M. *et al.* CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nature Cell Biol.* **11**, 172–182 (2009).
15. Thompson, B. A., Tremblay, V., Lin, G. & Bochar, D. A. CHD8 is an ATP-dependent chromatin remodeling factor that regulates β -catenin target genes. *Mol. Cell. Biol.* **28**, 3894–3904 (2008).
16. Batsukh, T. *et al.* CHD8 interacts with CHD7, a protein which is mutated in CHARGE syndrome. *Hum. Mol. Genet.* **19**, 2858–2866 (2010).
17. Betancur, C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res.* **1380**, 42–77 (2011).
18. Moller, R. S. *et al.* Truncation of the Down syndrome candidate gene *DYRK1A* in two unrelated patients with microcephaly. *Am. J. Hum. Genet.* **82**, 1165–1170 (2008).
19. Cooper, G. M. *et al.* A copy number variation morbidity map of developmental delay. *Nature Genet.* **43**, 838–846 (2011).
20. Buysse, K. *et al.* Delineation of a critical region on chromosome 18 for the del(18)(q12.2q21.1) syndrome. *Am. J. Med. Genet. A* **146A**, 1330–1334 (2008).
21. Hoischen, A. *et al.* *De novo* mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nature Genet.* **42**, 483–485 (2010).
22. Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38**, W214–W220 (2010).
23. Erten, S., Bebek, G., Ewing, R. & Koyutürk, M. DADA: Degree-aware algorithms for network-based disease gene prioritization. *BioData Mining* **4**, 19 (2011).
24. De Ferrari, G. V. & Moon, R. T. The ups and downs of Wnt signaling in prevalent neurological disorders. *Oncogene* **25**, 7545–7553 (2006).
25. Bedogni, F. *et al.* Tbr1 regulates regional and laminar identity of postmitotic neurons in developing neocortex. *Proc. Natl Acad. Sci. USA* **107**, 13129–13134 (2010).
26. Turner, E. H., Lee, C., Ng, S. B., Nickerson, D. A. & Shendure, J. Massively parallel exon capture and library-free resequencing across 16 genomes. *Nature Methods* **6**, 315–316 (2009).
27. Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).
28. Sakai, Y. *et al.* Protein interactome reveals converging molecular pathways among autism disorders. *Sci. Transl. Med.* **3**, 86ra49 (2011).
29. Gilman, S. R. *et al.* Rare *de novo* variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron* **70**, 898–907 (2011).
30. Ille, F. & Sommer, L. Wnt signaling: multiple functions in neural development. *Cell. Mol. Life Sci.* **62**, 1100–1108 (2005).
31. Tedeschi, A. & Di Giovanni, S. The non-apoptotic role of p53 in neuronal biology: enlightening the dark side of the moon. *EMBO Rep.* **10**, 576–583 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank and recognize the following ongoing studies that produced and provided exome variant calls for comparison: NHLBI Lung Cohort Sequencing Project (HL 1029230), NHLBI WHI Sequencing Project (HL 102924), NIEHS SNPs (HHSN273200800010C), NHLBI/NHGRI SeattleSeq (HL 094976), and the Northwest Genomics Center (HL 102926). We are grateful to all of the families at the participating Simons Simplex Collection (SSC) sites, as well as the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, E. Hanson, D. Grice, A. Klin, R. Kochev, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren and E. Wijsman). We also acknowledge M. State and the Simons Simplex Collection Genetics Consortium for providing Illumina genotyping data, T. Lehner and the Autism Sequencing Consortium for providing an opportunity for pre-publication data exchange among the participating groups. We appreciate obtaining access to phenotypic data on SFARI Base. This work was supported by the Simons Foundation Autism Research Initiative (SFARI 137578 and 191889; E.E.E., J.S. and R.B.) and NIH HD065285 (E.E.E. and J.S.). E.B. is an Alfred P. Sloan Research Fellow. E.E.E. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions E.E.E., J.S. and B.J.O. designed the study and drafted the manuscript. E.E.E. and J.S. supervised the study. R.B., B.R. and B.J.O. analysed the clinical information. R.B., L.V., S.G., E.K., N.K. and B.P.C. contributed to the manuscript. S.G., N.K., B.P.C., A.K., C.B., M.M. and L.V. generated and analysed CNV data. B.J.O. and L.V. performed MIP resequencing and mutation validations. I.B.S., E.H.T., B.J.O. and J.S. developed MIP protocol and analysis. B.V. and J.M.A. generated loci-specific mutation rate estimates. R.L. and E.B. performed PPI network analysis and simulations. E.K. performed DADA analysis. C.L. performed Illumina sequencing. J.D.S., I.B.S., E.H.T. and C.L. analysed sequence data. B.P.C. performed IPA analysis. B.J.O., E.K. and N.K. developed the *de novo* analysis pipelines and analysed sequence data. D.A.N., M.J.R., J.D.S. and E.H.T. supervised exome sequencing and primary analysis.

Author Information Access to the raw sequence reads can be found at the NCBI database of Genotypes and Phenotypes (dbGaP) and National Database for Autism Research under accession numbers phs000482.v1.p1 and NDARCOL0001878, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.E.E. (eee@gs.washington.edu) or J.S. (shendure@uw.edu).

METHODS

Exome capture, alignments and base-calling. Exomes for the 189 trios (and 31 unaffected siblings) were captured with NimbleGen EZ Exome V2.0. Final libraries were then sequenced on either an Illumina GAIIX (paired- or single-end 76-bp reads) or HiSeq2000 (paired- or single-end 50-bp reads). Reads were mapped to a custom GRCh37/hg19 build using BWA 0.5.6 (ref. 32). Read qualities were recalibrated using GATK Table Recalibration 1.0.2905 (ref. 33). Picard-tools 1.14 was used to flag duplicate reads (<http://picard.sourceforge.net/>). GATK IndelRealigner 1.0.2905 was used to realign reads around insertion/deletion (indel) sites. Genotypes were generated with GATK Unified Genotyper³³ with FILTER = "QUAL ≤ 50.0 || AB ≥ 0.75 || HRun > 3 || QD < 5.0" and in parallel with the SAMtools pipeline as described previously⁴. Only positions with at least eightfold coverage were considered. All pilot sibling exomes were captured and analysed as described previously⁴. Predicted *de novo* events were called and compared against a set of 946 other exomes to remove recurrent artefacts and likely undercalled sites. Indels were also called with the GATK Unified Genotyper and SAMtools and filtered to those with at least 25% of reads showing a variant at a minimum depth of 8×. Mutations were phased using molecular cloning of PCR fragments, read-pair information, linked informative SNPs, and obligate carrier status. To identify rare private variants (singleton), the full variant list was compared against a larger set of 1,779 other exomes. Predicted *de novo* indels were also filtered against this larger set.

Sanger validations. All reported *de novo* events (exome or MIP capture) were validated by designing primers with BatchPrimer3 followed by PCR amplification and Sanger sequencing. We performed PCR reactions using 10 ng of DNA from father, mother, unaffected sibling (when available), and proband and performed Sanger capillary sequencing of the PCR product using forward and reverse primers. In some cases, one direction could not be assessed due to the presence of repeat elements or indels in close proximity to the mutation event.

Mutation candidate gene analysis. We examined whether each non-synonymous or CNV *de novo* event may be contributing to the aetiology of ASD by evaluating the likelihood deleteriousness of the change (GERP, Grantham score) and intersecting with known syndromic and non-syndromic candidate genes, CNV morbidity maps, and information in OMIM and PubMed. Mutations were considered severe if they were truncating, missense with Grantham score ≥ 50 and GERP score ≥ 3 or only Grantham score ≥ 85, or deleted a highly conserved amino acid. For genes that had not previously been implicated in ASD, we gave priority to those with structural similarities to known candidate or strong evidence of neural function or development.

Exome read-depth CNV discovery. To find CNVs using exome read-depth data, we first mapped sequenced reads to the hg19 exome using the mrsFAST aligner³⁴. Next, we applied a novel method (N.K. *et al.*, manuscript in preparation), which uses normalized RPKM values³⁵ of the ~194,000 captured exons/sequences, subsequent population normalization using 366 exomes from the Exome Sequencing Project and singular value decomposition to remove systematic bias present within exome capture reactions. Rare CNVs were detected using a threshold cutoff of the normalized RPKM values, and we required at least three exons above our threshold in order to make a call. We made a total of 1,077 deletion or duplication calls in 366 individuals (range 0–14, median = 3, mean = 2.94).

CNV detection using array CGH. A custom-targeted 2 × 400K Agilent chip with median probe spacing of 500 bp in the genomic hotspots flanked by segmental duplications or Alu repeats and probe spacing of 14 kb in the genomic backbone was designed. All experiments were performed according to the manufacturer's instructions using NA12878 as the female reference and NA18507 as the male reference (Coriell). Data analysis was performed following feature extraction using DNA analytics with ADM-2 setting. All CNV calls were visually inspected in the UCSC Genome Browser. CNV calls from probands were then intersected with those from parents and also with 377 controls recruited through NIMH Genetics Initiative^{36,37} and ClinSeq cohort³⁸ analysed on the same microarray platform. The NIMH set of controls were ascertained by the NIMH Genetics Initiative³⁶ through an online self-report based on the Composite International Diagnostic Instrument Short-Form (CIDI-SF)³⁷. Those who did not meet DSM-IV criteria for major depression, denied a history of bipolar disorder or psychosis, and reported exclusively European origins were included^{39,40}. Samples from the ClinSeq cohort were selected from a population representing a spectrum of atherosclerotic heart disease³⁸. *De novo* and inherited potential pathogenic CNVs were selected only if they intersected with RefSeq coding sequence and allowing for a frequency of <1% in the controls and <50% segmental duplication content.

Illumina array CNV calling. CNV calling was performed in hg18 as described previously⁴¹, using an HMM that incorporates both allele frequencies (BAF) and total intensity values (logR). In total, we generated CNV calls for 841 probands, 1,651 parents and 793 siblings including the samples reported recently². Of the 122

families selected for CNV comparisons in this study, calls were generated for 107 probands. Of these, both parents were profiled for 101 families and one parent was profiled for the remaining six families. In addition, at least one sibling was profiled for 99 of these families.

Independent of array CGH detection, to identify putatively pathogenic CNVs, we first compared our data to 2,090 control samples derived from the Wellcome Trust Case Control Consortium (WTCCC) National Blood Services Cohort^{19,42} and filtered all CNVs present in 1% (20) of WTCCC2 controls or 1% (16) of parents by 50% reciprocal overlap with matching copy number status. In addition, similar to the filtering criteria used for array CGH detection, we selected only CNVs that contained less than 50% segmental duplication and intersected with RefSeq coding sequence. To select putative *de novo* CNVs, we further required the CNV not to be present in family-matched parents and siblings. Additionally, we filtered CNVs present in >0.1% (2) of the full 1,651 parent set. To select potential, rare inherited events, we required the CNV be detected in a matched parent or sibling. Finally, we filtered the genes inside each CNV under the same criteria (to account for smaller or larger CNPs) and removed CNVs with no remaining genes.

CNV cross validation. High-confidence, cross-validated *de novo* and inherited CNVs were selected by identifying events detected by at least two of three methodologies. To account for the variable breakpoint definitions in array CGH, SNP arrays, and exome copy number profiles, we aligned the CNVs by at least one overlapping gene ID and reported each CNV region by its maximal outer boundaries. This identified six *de novo* and 70 rare inherited events for further study (Supplementary Table 6).

Ingenuity pathway analysis. Ingenuity pathway analysis (IPA) was performed to identify potential functional enrichments within both our PPI (49 genes) and overall set of 126 genes. RefSeq reference gene list was used as a background list for all analysis. To confirm our results pertaining to *CTNNB1* upstream enrichment, we simulated 10,000 random populations of 209 individuals using Poisson priors for each gene based on their estimated mutation rates (see below), with a global correction factor resulting in selecting a mean of 126 genes per population. We then used this simulation data to calculate the probability of observing eight direct upstream interactors of *CTNNB1* and determined that our data set is enriched for these genes with $P = 0.0030$.

Estimating locus-specific mutation rates. Human–chimpanzee alignments were downloaded from the UCSC Genome Browser (reference versions GRCh37 and panTro2, <http://hgdownload.cse.ucsc.edu/goldenPath/hg19/vsPanTro2/syntenicNet/>). The more conservative syntenicNet alignments were used (details in <http://hgdownload.cse.ucsc.edu/goldenPath/hg19/vsPanTro2/README.txt>). Gene definitions were downloaded from the UCSC Table Browser, from the RefSeq Genes track, and the refFlat table. Exons were extended by 2 bp, and overlapping exons were merged using BEDTools. Non-exonic sequence was not considered. For each gene, we extracted: (1) d = the number of differences between chimpanzee and human; and (2) n = the number of bases aligned. We assumed a divergence time between human and chimpanzee of 12 million years (Myr) and an average generation time of 25 years. We then calculated gene-specific mutation rates per site per generation: $r = (d/n)/(12 \text{ Myr}/25 \text{ years/generation})$. We calculated the probability of observing X + events using the Poisson distribution defined by the number of chromosomes screened and the size of the coding region, including actual splice bases.

Network simulation and null model estimation. To generate a null distribution of gene mutations, *de novo* mutation rates were estimated from human–chimp mutation rates. A pseudocount of 2.0833×10^{-6} (the smallest calculated in the gene set) was applied to any exon with a mutation rate of zero. To create null gene sets, genes were drawn uniformly from this background distribution. Human protein–protein interaction data were collected from GeneMANIA²² on 29 August 2011. Only direct physical interactions from the *Homo sapiens* database were considered. The list comprises approximately 1.5 million physical interactions, gathered from 150 studies. A protein interaction network was created from each experimental and null gene set by drawing edges between genes with physical interactions reported in the GeneMANIA database. Qualitatively similar results were achieved by including only interactions supported by multiple independent data sources. For each network, clustering coefficient, centralization, average shortest path length, density, and heterogeneity were determined using Cytoscape⁴³ and Network Analyzer⁴⁴. Duplicate- and self-interactions were not considered in calculating network statistics.

Disease gene prioritization based on PPI networks. We applied degree-aware algorithms to rank a set of candidate genes with respect to a set of products of genes associated with ASD using human PPI networks. We used the integrated human PPI network data collected from GeneMANIA²² on 29 August 2011. The PPI network contains 12,007 proteins with ~1.5 million direct physical interactions associated with a reliability score. We obtain the seed proteins for the ASD from the list of ref. 17. For the candidate set we used 126 gene products from the severe

disruptive *de novo* events from the pilot autism project⁴ and the current study. Given the GeneMANIA PPI network and Betancur seed gene product list, we used DADA²³ for ranking the candidate genes. We emphasize that this ranking is not implying causality but rather relatedness to genes previously and independently associated with ASD. For testing the significance of this ranking, we rank all the gene products except the seed set using the same algorithm. On the basis of the ranking result, we applied a Mann–Whitney *U* rank sum test (one-tailed) on the candidate set compared to all the other genes.

MIP protocol. Each of 1,703 autism probands from the SSC collection and 744 controls from the NIMH collection was subjected to MIP-based multiplex capture of the six genes: *SCN1A*, *GRIN2B*, *GRIN2A*, *LAMC3*, *FOXP1* and *FOXP2*. For each library, 50 ng of DNA was used. Individually synthesized 70 mer MIPs ($n = 355$) were pooled and 5' phosphorylated with T4 PNK (NEB). Hybridization with MIPs, gap filling and ligation were performed in one step for 45–48 h at 60 °C, followed by an exonuclease treatment of 30 min at 37 °C, similar to ref. 45, with modifications for reduced MIP number (B.J.O. *et al.*, manuscript in preparation). Amplification of the library was performed by PCR using different barcoded primers for each library. Then barcoded libraries were pooled, purified using Agencourt AMPure XP and one lane of 101-bp paired-end reads was generated for each mega-pool (~384) on an Illumina HiSeq 2000 according to manufacturer's instructions. Raw reads were mapped to the genome as in ref. 4. MIP targeting arms were then removed and variants called using SAMtools⁴. A 25-fold coverage, with AB allele ratio <0.7 , and quality 30 threshold was used for high-confident variant calling. Private (possible *de novo*) variants were identified by filtering against 1,779 other exomes. The parents of children with disruptive rare variants were then captured. Variants not seen or with low coverage in the parents were validated by Sanger capillary-based fluorescent sequencing. No truncating variants of *GRIN2B* were observed in the MIP sequenced controls or the Exome Variant Server ESP2500 release (NHLBI Exome Sequencing Project (ESP), Seattle, Washington, <http://evs.gs.washington.edu/EVS/>).

Estimating the number of autism loci. The gene-level specificity of exome sequencing enables the estimation of the number of recurrently mutated genes implicated in the genetic aetiology of sporadic ASD. This question can be reformulated as the 'unseen species problem' (see ref. 46 for review and ref. 2 for application to *de novo* CNVs discovered in autism), where genes with severe *de novo* events in probands are considered 'observed species', and binned by their frequency of appearance (that is, singletons, doubletons, etc.). We estimated the total number of genes implicated in autism (the total number of species) using several different estimators (implemented in the R package SPECIES, <http://www.jstatsoft.org/>), as well as the formula provided in ref. 2. This estimate depends on the number of singletons and twin pairs of genes observed in probands, as well as the fraction of *de novo* events believed to be pathogenic for autism, that is, single, disruptive events that can cause autism on their own. We assumed that both of our

recurrent severe *de novo* events (affecting *CHD8* and *NTNG1*) were pathogenic; these compose the entire set of twin pairs. The number of singletons is based on the estimated a priori fraction of the observed events that are pathogenic for autism. Across this sliding scale, the estimated number of loci is plotted in Supplementary Fig. 13. For example, using the estimator from ref. 47, if 20–50% of our *de novo* severe events are considered pathogenic, exome sequencing of a large number of additional samples would reveal between 182 and 992 pathogenic genes harbouring coding *de novo* point mutations (Supplementary Fig. 13); if all the observed severe *de novo* events in our experiment are included as pathogenic singletons, the number of implicated loci increases to more than 3,000.

32. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
33. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genet.* **43** (2011).
34. Hach, F. *et al.* mrsFAST: a cache-oblivious algorithm for short-read mapping. *Nature Methods* **7**, 576–577 (2010).
35. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).
36. Moldin, S. O. NIMH Human Genetics Initiative: 2003 update. *Am. J. Psychiatry* **160**, 621–622 (2003).
37. Kessler, R. C. & Ustun, T. B. The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int. J. Methods Psychiatr. Res.* **13**, 93–121 (2004).
38. Biasecker, L. G. *et al.* The ClinSeq Project: piloting large-scale genome sequencing for research in genomic medicine. *Genome Res.* **19**, 1665–1674 (2009).
39. Talati, A., Fyer, A. J. & Weissman, M. M. A comparison between screened NIMH and clinically interviewed control samples on neuroticism and extraversion. *Mol. Psychiatry* **13**, 122–130 (2008).
40. Baum, A. E. *et al.* A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol. Psychiatry* **13**, 197–207 (2008).
41. Itsara, A. *et al.* Population analysis of large copy number variants and hotspots of human genetic disease. *Am. J. Hum. Genet.* **84**, 148–161 (2009).
42. Craddock, N. *et al.* Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* **464**, 713–720 (2010).
43. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
44. Assenov, Y., Ramirez, F., Schelhorn, S. E., Lengauer, T. & Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **24**, 282–284 (2008).
45. Mamanova, L. *et al.* Target-enrichment strategies for next-generation sequencing. *Nature Methods* **7**, 111–118 (2010).
46. Bunge, J. & Fitzpatrick, M. Estimating the number of species - a Review. *J. Am. Stat. Assoc.* **88**, 364–373 (1993).
47. Chao, A. & Lee, S. M. Estimating the number of classes via sample coverage. *J. Am. Stat. Assoc.* **87**, 210–217 (1992).

Unexpected features of *Drosophila* circadian behavioural rhythms under natural conditions

Stefano Vanin^{1*†}, Supriya Bhutani^{2*†}, Stefano Montelli^{1†}, Pamela Menegazzi^{1†}, Edward W. Green², Mirko Pegoraro^{1,2}, Federica Sandrelli¹, Rodolfo Costa¹ & Charalambos P. Kyriacou²

Circadian clocks have evolved to synchronize physiology, metabolism and behaviour to the 24-h geophysical cycles of the Earth¹. *Drosophila melanogaster*'s rhythmic locomotor behaviour provides the main phenotype for the identification of higher eukaryotic clock genes^{2,3}. Under laboratory light–dark cycles, flies show enhanced activity before lights on and off signals^{4,5}, and these anticipatory responses have defined the neuronal sites of the corresponding morning (M) and evening (E) oscillators^{6,7}. However, the natural environment provides much richer cycling environmental stimuli than the laboratory, so we sought to examine fly locomotor rhythms in the wild. Here we show that several key laboratory-based assumptions about circadian behaviour are not supported by natural observations. These include the anticipation of light transitions, the midday 'siesta', the fly's crepuscular activity, its nocturnal behaviour under moonlight, and the dominance of light stimuli over temperature. We also observe a third major locomotor component in addition to M and E, which we term 'A' (afternoon). Furthermore, we show that these natural rhythm phenotypes can be observed in the laboratory by using realistic temperature and light cycle simulations. Our results suggest that a comprehensive re-examination of circadian behaviour and its molecular readouts under simulated natural conditions will provide a more authentic interpretation of the adaptive significance of this important rhythmic phenotype. Such studies should also help to clarify the underlying molecular and neuroanatomical substrates of the clock under natural protocols.

Drosophila melanogaster provides a prominent model system in higher eukaryotes for studying the molecular and neurogenetic basis of circadian behavioural rhythms². The fly's 24-h rhythmic locomotor activity has morning (M) and evening (E) components, interpreted in numerous laboratory studies to reflect the fly's anticipation of regular changes in light–dark transitions^{4,5} (Fig. 1a). Neurogenetic dissection has revealed discrete M and E neuronal clusters that determine these circadian landmarks^{6,7}. We wondered how the complex naturally cycling geophysical environment might affect these behavioural patterns, so we investigated fly locomotor rhythms in wild-type and clock mutant strains under natural conditions, by placing activity monitors outdoors for three seasons (2007–2009), from April to November in Leicester, UK (latitude 52° 38' N), and Treviso, Italy (45° 65' N) (Supplementary Fig. 1).

Generally, and throughout the seasons, clock mutants show very high levels of rhythmicity (85–100%) that reflect the cycling environment driving (entraining) their behaviour, except *Clk^{rk}*, which shows decreased rhythmicity (40%) at lower temperatures (Supplementary Table 1). Figure 1b–d illustrates locomotor profiles of a natural wild-type strain (WTALA) and clock mutants from summer observations in Treviso. Contrary to laboratory-based results, at warmer temperatures the 'siesta' (decreased activity during the hottest parts of the day;

Fig. 1a)^{8–10} was not observed. Instead, and in addition to the well-established M and E locomotor peaks, a prominent mid-afternoon component, which we term 'A', was present in more than 50% of wild-type flies at mean daytime temperatures of 20 °C, rising to about 100% at 27 °C (Fig. 1c). Laboratory studies in light–dark cycles at constant high temperatures of 29 °C (refs 8, 9), or in rectangular light and temperature cycles, do not reveal A components (Fig. 1a). This previously unobserved behaviour was seen in all genotypes at warmer temperatures (Fig. 1d and Supplementary Fig. 2) and may represent a stress/escape response that is phase-locked to the rise and fall in temperature. Alternatively, but not exclusively, A may reflect an environmentally modulated circadian phenotype. We observed that the A component was significantly advanced by up to 3 h in short-period *per^S* mutants in comparison with other *per* genotypes; *per⁰¹* was also advanced compared with *per^L* and the wild type, and the A offset was significantly delayed in *per^L* compared with the other genotypes (Fig. 1e and Supplementary Fig. 2), suggesting that A is modulated by the clock and is not a simple escape reflex.

Under laboratory 12 h light–12 h dark cycles, wild-type flies anticipate lights-on by increasing their locomotor activity about 2 h before the lights-on signal, in contrast with arrhythmic mutants such as *per⁰¹*, which do not anticipate the signal^{6,7}. Consequently, in nature, if wild-type flies anticipate a dawn-related geophysical transition, clock mutants would be expected to 'not anticipate' and delay their activity accordingly. In nature, astronomical, nautical and civil twilights, defined by the angles of the Sun below the horizon (18°, 12° and 6°, reflecting illumination of approximately 0.001, 0.01 and about 1 lx, respectively), accompany the dawn–dusk transitions¹¹. We therefore used nautical twilight, which occurs roughly 1.5–2 h before dawn, as a convenient environmental marker by which to compare the onset of morning activity (M_{onset}) among genotypes. Seasonally, the average daily M_{onset} for WTALA flies in Leicester and Treviso reveals that the two locations differ in their overall profiles (Supplementary Fig. 3a). However, when the values are replotted against average night temperature, this difference evaporates, revealing a similar inverse relationship between M_{onset} and temperature in three wild-type strains (Fig. 2a–c), clock mutants and Gal4-driven genotypes (Fig. 2d–j and Supplementary Fig. 4). The exception is *tim⁰¹*, for which M_{onset} is only marginally temperature-dependent (Fig. 2d) and is rescued by transformation with the *tim⁺* gene, and in *per⁰¹;tim⁰¹* double mutants (Fig. 2e, f). Ablating the Pigment Dispersing Factor (PDF)-expressing (Morning) neurons with *UAS-hidUAS-rpr*, or using the *UAS-CycA* dominant-negative to stop the clock in these cells, failed to disrupt the relationship between M_{onset} and temperature (Supplementary Fig. 4).

We investigated whether clock mutants 'anticipate' changes in twilight differently from the wild type, mirroring laboratory studies^{6,7}. We used an analysis of covariance (ANCOVA; temperature as covariate) to

¹Department of Biology, University of Padova, Padova 35131, Italy. ²Department of Genetics, University of Leicester, LE1 7RH, UK. [†]Present addresses: Department of Chemical and Biological Sciences, School of Applied Science, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK (S.V.); Department of Cellular and Molecular Neurosciences, National Brain Research Centre, Manesar NH8, Haryana, 122050, India (S.B.); Department of Experimental Veterinary Science, University of Padova, Viale dell'Università 16, 35020 Legnaro (Agripolis) Padova, Italy (S.M.); Lehrstuhl für Neurobiologie und Genetik, Universität Würzburg, Biozentrum, Am Hubland, 97074 Würzburg, Germany (P.M.).

*These authors contributed equally to this work.

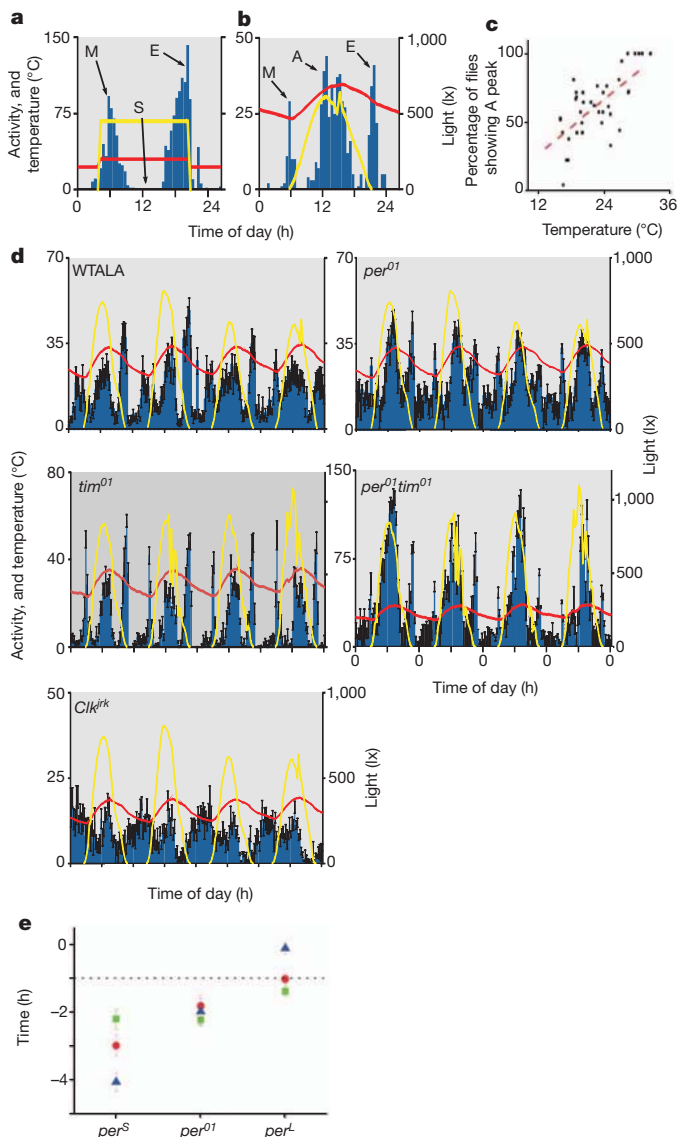


Figure 1 | The afternoon (A) component of locomotor activity. **a**, An individual male's (WTALA) locomotor activity in laboratory conditions with 16 h light–8 h dark and 22 °C in the dark phase and 30 °C in the light. M, morning; S, siesta; E, evening (yellow, light intensity; red, temperature). **b**, An individual male's (WTALA) locomotor activity on a single summer's day (yellow, light intensity; red, temperature; day temperatures: average 28.7 °C, maximum 34.4 °C, minimum 23.4 °C). M, morning; A, afternoon; E, evening. Time 0 = midnight (00:00). **c**, Afternoon peak is associated with temperature ($R^2 = 0.34$, $F_{1,32} = 17.7$; $P < 0.0002$). **d**, Averaged locomotor plots (means \pm s.e.m.) for WTALA and four arrhythmic mutants (see Supplementary Fig. 2 for other mutants). WTALA ($n = 19$), per^{01} ($n = 12$) and Clk^{irk} ($n = 27$, with 3 arrhythmic flies) were monitored together (average temperature over experiment 28.0 °C, maximum 34.4 °C, minimum 21.1 °C), whereas tim^{01} ($n = 17$) and the double mutant $per^{01};tim^{01}$ ($n = 12$) were monitored at a different time (average temperature 29.3 °C, maximum 35.9 °C, minimum 23.1 °C). Red, temperature; yellow, light intensity. **e**, Afternoon component in per mutants compared with wild-type ALA (WTALA). Times of A_{onset} (green squares), A_{peak} (red circles) and A_{offset} (blue triangles) of WTALA were subtracted from the corresponding values for the mutants because not all genotypes were run in the same experiment. Negative values indicate an earlier phenotype than WTALA. ANOVAs for all parameters gave significant genotype effects ($F_{2,151} = 4.8$ (onset), 8.9 (peak) and 29.8 (offset); all $P < 0.001$). In *post hoc* tests, per^S was significantly earlier in all parameters than per^L , and significantly earlier than per^{01} in peak and offset. per^L was significantly delayed compared with per^{01} in all parameters.

inspect the average time from nautical twilight that each clock variant began morning activity, but only Pdf^0 , which was slightly delayed,

fell significantly outside the range of the three wild types (Fig. 2i and Supplementary Table 2a). These results suggest that clock genotypes (except tim^{01}) generally anticipate geophysical variables equally. Alternatively, if M_{onset} activities are simply temperature-sensitive responses to dim light levels, they should be delayed on the following morning if the activity monitors are covered up after the previous sunset. When we performed this experiment, significant delays of about 1 h were observed for all genotypes (except tim^{01}), with M_{onset} now tracking the morning rise in temperature (Fig. 2k, l, Supplementary Fig. 5 and Supplementary Table 2b). The attenuated tim^{01} M_{onset} twilight-modulated temperature responses (Fig. 2d and Supplementary Fig. 5a) recall the mutant's disrupted diapause, a photoperiodic phenotype also requiring both light and temperature input¹². We conclude that M_{onset} is a twilight-dependent temperature response with little evidence for circadian regulation.

We next examined evening onset (E_{onset}), which is seasonally delayed relative to the photoperiod during the warmer summer months (Supplementary Fig. 3b). We replotted E_{onset} against average daytime temperature, relative to the time of maximum daily temperature (T_{max} ; Fig. 3 and Supplementary Fig. 6). All genotypes showed E_{onset} activities that began successively later as mean daily temperatures rose above 20 °C (Fig. 3 and Supplementary Fig. 6). At less than 20 °C, the response was stable in relation to T_{max} , with marginally significant negative relationships for WTALA, HU and Clk^{irk} strains (Supplementary Table 3a) and with E_{onset} generally occurring within the 2.25 ± 0.17 h (mean \pm s.e.m.) that defined the time between the maximum daily light intensity and the subsequent T_{max} (Fig. 3a–i). When analysed separately for temperatures below and above 20 °C, per^S showed the earliest E_{onset} activities, followed by per^{01} and tim^{01} ; per^L had the latest (Supplementary Table 3b, c). This is reflected in Fig. 3, where the data points fall either mostly above (later, per^L) or below (earlier, per^S , per^{01} , tim^{01}) the upper dotted line representing T_{max} . These results mirror laboratory findings in which per^S and per^L modulate the timing of evening behaviour¹³; in addition, however, the unexpected observation of E components in arrhythmic mutants, and their phase advance, indicates an underlying residual short-period rhythmicity^{5,14–16}.

In laboratory wild-type flies, the length of the temperature-sensitive afternoon 'siesta' is mediated by the thermal regulation of per 3' splicing^{8–10}. E_{onset} defines the end of the siesta, so we studied whether the per splicing readout correlated with the E_{onset} curve. Levels of per 3' splicing in fly heads throughout the seasons for WTALA flies (Fig. 3j) were linearly related to temperature over the entire range 7–30 °C, so it is unlikely that splicing contributed significantly to E_{onset} at temperatures below 20 °C (Supplementary Fig. 6F). Indeed, because the laboratory siesta is mediated by the dynamics of the upswing of the clock protein PERIOD (PER)⁸, the fact that per^{01} -null mutants showed a similar E_{onset} –temperature relationship to that of the wild type also argues against a significant role for per 3' splicing in this natural phenotype. A similar conclusion using splice-locked per transformants was reached in laboratory experiments using artificial temperature and light cycles¹⁷.

Drosophila rhythms in the laboratory are often described as 'crepuscular' because of the distribution of M and E components at dawn and dusk¹⁸. We therefore quantified the activity falling in the morning and evening twilights (astronomical, nautical and civil) as a proportion of all activity falling between the beginning of morning and the end of evening astronomical twilights. Using data from Italian experiments, we observed that even at the warmest temperatures when M_{onset} and E_{onset} are pushed towards the twilights (Figs 1–3), the proportion of activity falling within the twilights was less than 25% and was reduced still further at cooler temperatures (Fig. 3k). Thus, in nature the major proportion of activity, even at high temperatures, falls outside the twilights, so flies are diurnal rather than crepuscular.

Related laboratory experiments have also suggested that locomotor activity becomes predominantly crepuscular and nocturnal under

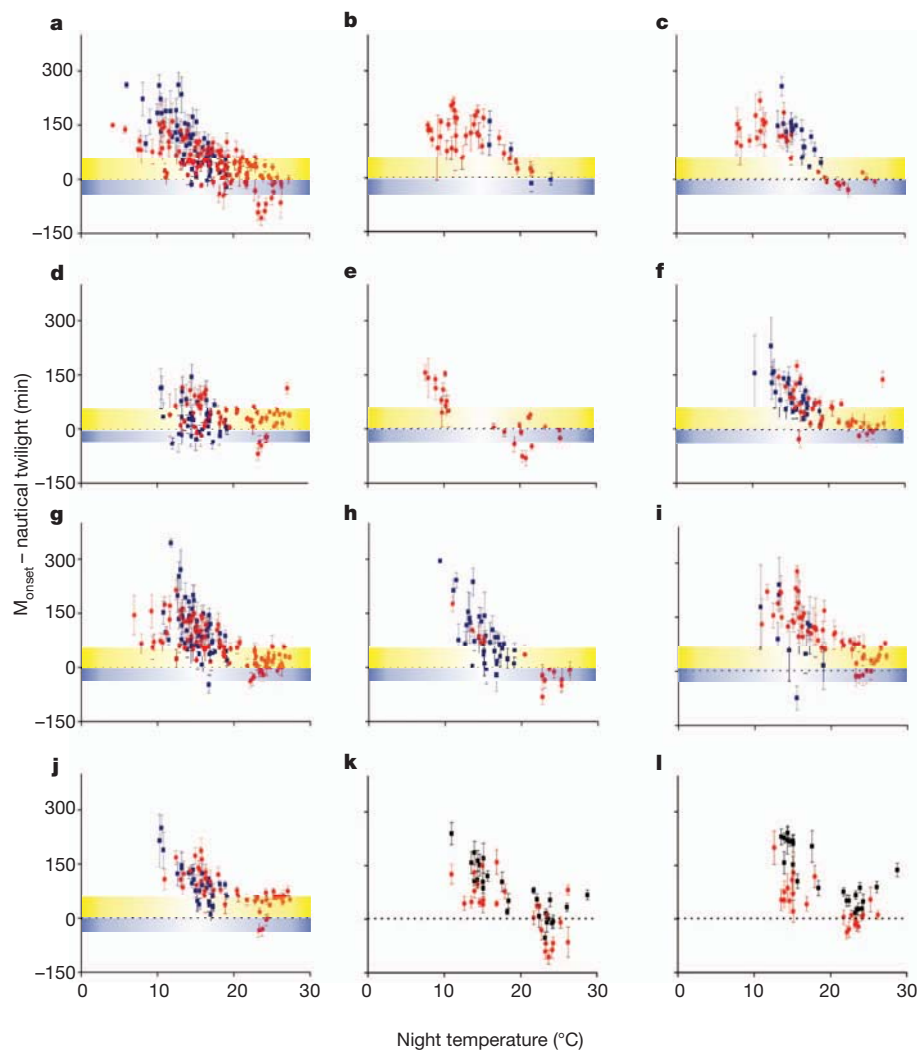


Figure 2 | The morning onset (M_{onset}) of locomotor activity is dependent on temperature and twilight. **a–j**, Time of M_{onset} in a single day compared with nautical twilight (dotted line, time 0). Blue and yellow areas represent the approximate time between astronomical and nautical twilight and between nautical twilight and sunrise, respectively (both calculated as seasonal averages for Treviso). Each point represents the mean of 4–24 flies (median 14); error bars represent s.e.m. Blue squares, Leicester; red circles, Treviso. **a–c**, Wild

types: WTALA (Alto Adige) (**a**), Canton S (**b**) and HU (Houten) (**c**). **d–j**, Clock mutants: tim^{01} (**d**), $tim^{01};P[tim^{+}]$ (**e**), $per^{01};tim^{01}$ (**f**), per^{01} (**g**), Clk^{jrk} (**h**), Pdf^0 (**i**) and cry^{01} (**j**) (Supplementary Fig. 4). The slope for tim^{01} is significantly different from that for all three wild types (Supplementary Table 2a). **k, l**, M_{onset} of wild type (**k**) and per^{01} mutant (**l**) before (red) and after (black) activity monitors were covered at sunset and maintained until mid-afternoon (Supplementary Fig. 5 and Supplementary Table 2b).

simulated moonlight and this correlates with delayed expression of the clock proteins PER and TIMELESS (TIM) in the fifth PDF-negative sLNv neuron^{19,20}. We calculated the proportion of WTALA nocturnal activity during new moons and immediately before and after full moons. Figure 4a reveals no significant differences in nocturnal activity between the two conditions, even when we varied the exposure to moonlight from 0.03 to 0.33 lx (Supplementary Fig. 7). When we analysed PER in adult brains by immunocytochemistry under conditions of full moon and no moon, consistent with the behaviour, PER expression in the fifth sLNv (Fig. 4b) was not significantly delayed in moonlight in comparison with the other lateral clock neurons (s-LNvs, l-LNvs and LNDs). The dorsal DN1 and DN2 neurons showed significantly earlier PER peaks than the other neurons (Fig. 4b and Supplementary Fig. 8), indicative of a faster underlying oscillation²¹. These results suggest that natural temperature and light cycles may counteract the delayed PER expression of the fifth sLNv observed under simulated moonlight.

Finally, we attempted to reproduce natural circadian phenotypes within the laboratory by simulating cycling temperature and light intensity. We simulated consecutive days of new moon (darkness at night), with astronomical twilight (0.005 lx) giving way gradually to

nautical twilight (0.05 lx), to civil twilight (1 lx) over a 2-h period at the beginning and (in reverse) at the end of the day, rising to a maximum intensity of 1,500 lx in the afternoon. We ran this at cycling day–night temperatures ranging from 20 to 30 °C (spring) or from 25 to 35 °C (summer), mimicking actual temperatures that we measured in two Treviso spring and summer days, respectively, when the photoperiod for both was roughly 16 h light–8 h dark. We observed that in the summer simulation, M, A and E components were prominently displayed by all genotypes (WTALA, per mutants, tim^{01} and Clk^{jrk}), although Clk^{jrk} showed a smaller E burst on most days (Fig. 4c and Supplementary Fig. 9). We confirmed that M_{onset} occurs between the simulated times of astronomical and nautical twilight, as in the wild at these temperatures, with all genotypes becoming active 0.2–0.7 h before simulated summer nautical twilight, but later (between 0.1 and 0.4 h) before simulated spring twilight (Supplementary Fig. 9c). As in the wild, M_{onset} is temperature-sensitive, with no significant differences between the wild type and either per^{01} or tim^{01} mutants in the timing on M_{onset} (although per^S and per^L were slightly earlier in their M_{onset} for both spring and summer simulations), and with tim^{01} showing a similar temperature independence to that in the wild (Supplementary Fig. 9c). We observed the circadian modulation of

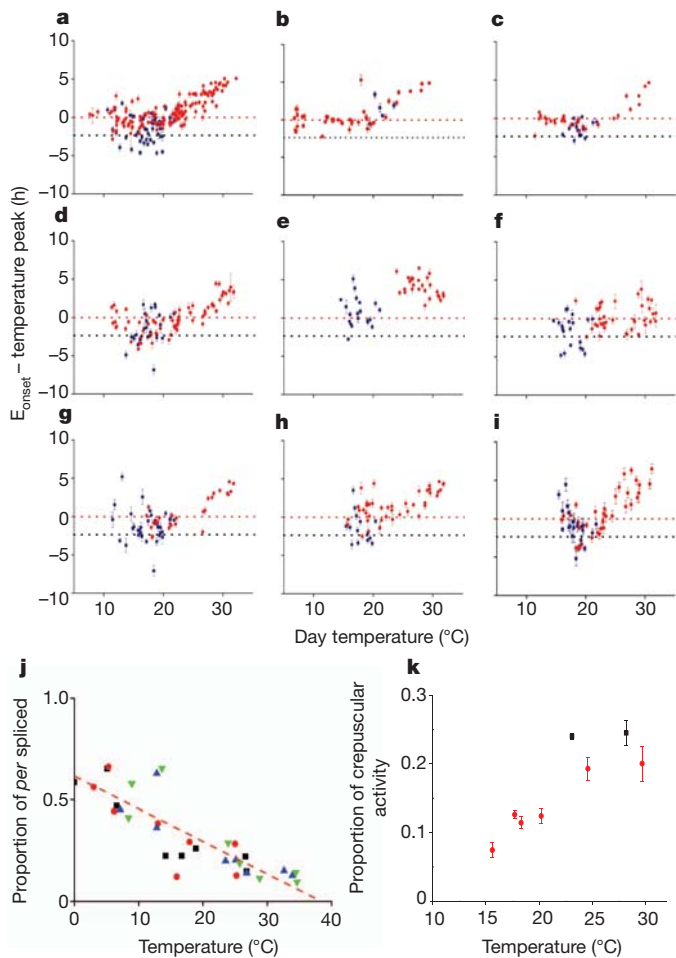


Figure 3 | Evening onset is temperature and clock modulated. **a–i**, Evening onset compared with daily maximum temperature (T_{\max} , 0). Dotted line below T_{\max} represents the average time that light peaked (about 2.25 h) before T_{\max} . Blue squares, Leicester; red circles, Treviso. **a–c**, Wild types: WTALA (Alto Adige) (a), Canton S (b) and HU (Houten) (c). **d–i**, Clock mutants: *per⁰¹* (d), *per^L* (e), *per^S* (f), *tim⁰¹* (g), *per⁰¹;tim⁰¹* (h) and *Clk^{rk}* (i) (see also Supplementary Fig. 6 and Supplementary Table 3). Most data points below 20 °C fall between the times of the light and temperature peaks. *per⁰¹*, *tim⁰¹* and *per^S* have earlier E_{onset} (ANCOVA: $P \approx 0$, 0.003 and 0.0007, respectively, versus WTALA) as shown by the lower scores on the y axis; *per^L* has later onset ($P = 0.0003$ versus wild-type). **j**, *per* splicing in nature in fly heads. *per^{spliced}* isoform compared with *per^{total}* (*per^{spliced}* + *per^{unspliced}*) against average daily temperature in WTALA strain. Heads collected at 00:00 (black squares), 03:00 (red circles), 12:00 (blue triangles) and 15:00 (green triangles). **k**, Locomotor activity is neither crepuscular in natural conditions nor in simulated natural conditions. Red circles, mean \pm s.e.m. of daily crepuscular activity in natural condition of groups of WTALA flies in Italy from single multi-day experiments; black squares, Italian spring and summer laboratory simulations (see Fig. 4c).

the E component in our simulations (Supplementary Fig. 9e) and whereas nearly all flies, irrespective of clock genotype, showed a prominent A component in the summer simulation, this proportion fell to 20–50% in the spring simulation, matching our results from nature and revealing the importance of temperature for the expression of A (Fig. 4c and Supplementary Fig. 9b, d). We also observed similar proportions of crepuscular activity for the spring and summer simulations to those in our natural observations (Fig. 3k). We did not, however, reproduce the advance in the A component of *per^S* compared with the wild type that we observed in nature (Supplementary Fig. 9d).

Our results from the wild reveal a diminished role for ‘anticipation’ of light–dark transitions partly because arrhythmic mutants also show M and E components, and also because the M component appears to be a response to changes in the environment. However, the newly

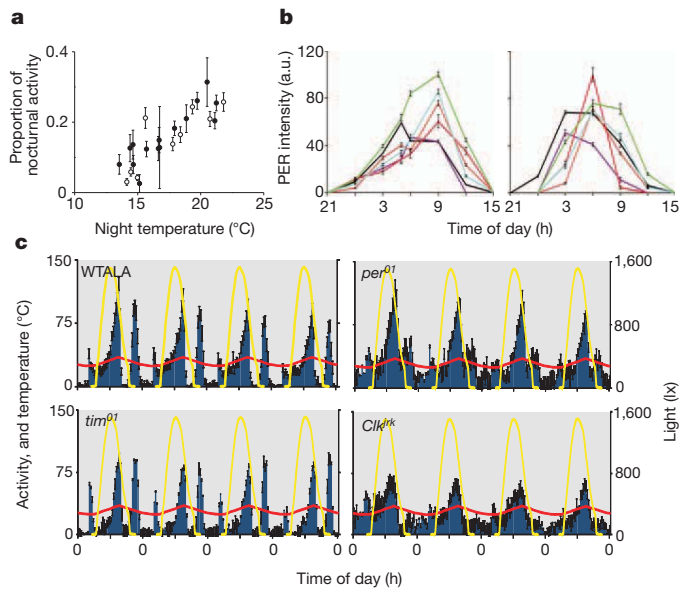


Figure 4 | Diurnal phenotypes of *D. melanogaster*. **a**, The proportion of nocturnal locomotor activity of WTALA males does not differ between days of full moon (open circles) and no moon (filled circles) (ANCOVA: $F_{1,22} = 0.04$, $P = 0.84$). Results are means \pm s.e.m. and are based on 4–24 flies (median 17). **b**, PER immunofluorescence intensity (mean \pm s.e.m., arbitrary units) in clock neurons in May 2008 (left) and September 2008 (right) (new and full moon, respectively, corresponding to filled and open circles in a). Red, fifth Pdf-null sLNv; brown, sLNvs; green, LNDs; turquoise, l-LNvs. DN1 neurons (black) and DN2 neurons (purple) are significantly advanced ($F_{5,498} = 13.71$, $P \approx 0$; see also Supplementary Fig. 8). There is a significant advance for all neurons in September in comparison with May ($F_{1,498} = 34.1$, $P \approx 0$) but no month \times neuron interaction ($F_{5,498} = 0.79$, $P = 0.56$). May, lights on at 05:16, lights off at 20:59, photoperiod 15 h 43 min; September, lights on at 06:20, lights off at 19:50, photoperiod 13 h 30 min. **c**, Simulation of Italian summer in the laboratory for the wild type and for arrhythmic mutants (16 h light–8 h dark, temperature range 25–35 °C; see also Supplementary Fig. 9). $n = 19$ –32 for each genotype at each temperature.

described A and the E components are clock-modulated, unlike M. This may be because M occurs during the rapid changes in illumination from quasi-darkness to bright light, when temperature is constant at its daily minimum, whereas A and E components occur when the light intensity is peaking or falling rapidly but temperature is either still rising or falling gradually. Such conflicting environmental signals may have recruited the clock to regulate appropriately adaptive A and E locomotor responses. Our results also show that temperature is the critical variable for predicting circadian behaviour¹⁷, because the large differences in photoperiod between our two locations (more than 3 h in midsummer) contribute little to the behavioural variance. In contrast, laboratory studies that place temperature and light cycles in different phases to each other²² reveal that light is the most important driver of circadian behaviour.

Although lateral and dorsal circadian neurons underlying M and E behaviour in light–dark cycles at constant temperature have been identified^{6,7}, subsets of dorsal neurons and lateral posterior neurons may be relevant for circadian temperature signalling^{22–26}. Perhaps in the wild it is these neurons that dominate the neuronal circuit. However, our observation that various ‘clockless’ genotypes can nevertheless show M and E (and A) components also suggests that the underlying neuronal substrates for this behavioural programme may not reside solely in the lateral and dorsal neurons but also elsewhere in the brain. Clock gene expression enhances and modulates circadian behaviour in the wild, but the possibility remains that a residual short-period rhythmicity resides within *per* and *tim* ‘arrhythmic’ mutants^{5,14–16} that can be amplified under natural conditions. Our results resonate with studies of laboratory mice²⁷ and hamsters²⁸,

which show them to behave quite differently in nature, where they are predominantly diurnal and not nocturnal as in the laboratory. Similarly, generally accepted adaptive and mechanistic explanations for fly circadian behaviour from laboratory experiments may require some revision if they are to account for rhythmicity in nature.

METHODS SUMMARY

Fly strains. Strains WTALA and HU are natural isolates collected as isofemale lines from northern Italy and The Netherlands¹². Canton-S is a standard laboratory strain. *per*⁰¹, *per*^S and *per*^L are congenic and Cantonized, but the genetic backgrounds of the other mutants used *per*⁰¹; *tim*⁰¹, *Clk*^{rk}, *Pdf*⁰, *cry*⁰ and Gal4-driven transgenes (drivers *tim-gal4* and *Pdf-gal4*), the apoptotic *UAS-hidUAS-rpr* and the dominant-negative *UAS-CycA* (ref. 29) are not equivalent, and for most are unknown.

Behaviour. Flies were raised at 25 °C in 12 h light–12 h dark cycles in the laboratory before males were placed outside at 2–3 days of age in Trikinetics activity monitors. Morning (*M*_{onset}) and afternoon or evening (*A*_{onset}, *E*_{onset}) onsets were determined operationally for each fly on each day of every experiment. Circadian rhythmicity was assessed by spectral analysis and autocorrelation as reported previously³⁰.

Moonlight measurements. Measurements of moonlight were made in Treviso and Leicester, using a LI-210SA Photometric Sensor (LI-COR) connected to a LI 1400 data logger.

Twilights. Twilights were obtained from the online database of the United States Naval Observatory (USNO) Astronomy Application Department¹¹.

Immunocytochemistry. The slides were examined under a Nikon 80i microscope with QICAM Fast Camera. Individual images were taken of planes at different depths to create a z-series for each brain. Brains were stained with primary antibodies against PDF and PER (the latter was a gift from R. Stanewsky).

RNA quantification. Flies were collected every 3 h for a 24-h time course in natural conditions, and fixed in liquid nitrogen. Heads were removed and *per*^{spliced}, *per*^{unspliced} forms and the control gene *Cbp20* (*Cap binding protein 20*) were amplified by PCR. PCR products were revealed on a 2% agarose gel and imaged, and final quantification was obtained with ImageJ software.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 July 2011; accepted 24 February 2012.

Published online 4 April 2012.

- Dodd, A. N. *et al.* Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **309**, 630–633 (2005).
- Hardin, P. E. Molecular genetic analysis of timekeeping in *Drosophila*. *Adv. Genet.* **74**, 141–173 (2011).
- Gallego, M. & Virshup, D. M. Post-translational modifications regulate the ticking of the circadian clock. *Nature Rev. Mol. Cell Biol.* **8**, 139–148 (2007).
- Helfrich-Forster, C. The locomotor activity rhythm of *Drosophila melanogaster* is controlled by a dual oscillator system. *J. Insect Physiol.* **47**, 877–887 (2001).
- Collins, B. H., Dissel, S., Gaten, E., Rosato, E. & Kyriacou, C. P. Disruption of Cryptochrome partially restores circadian rhythmicity to the arrhythmic period mutant of *Drosophila*. *Proc. Natl Acad. Sci. USA* **102**, 19021–19026 (2005).
- Grima, B., Chelot, E., Xia, R. & Rouyer, F. Morning and evening peaks of activity rely on different clock neurons of the *Drosophila* brain. *Nature* **431**, 869–873 (2004).
- Stoleru, D., Peng, Y., Agosto, J. & Rosbash, M. Coupled oscillators control morning and evening locomotor behaviour of *Drosophila*. *Nature* **431**, 862–868 (2004).
- Majercak, J., Sidote, D., Hardin, P. E. & Edery, I. How a circadian clock adapts to seasonal decreases in temperature and day length. *Neuron* **24**, 219–230 (1999).
- Collins, B. H., Rosato, E. & Kyriacou, C. P. Seasonal behavior in *Drosophila melanogaster* requires the photoreceptors, the circadian clock, and phospholipase C. *Proc. Natl Acad. Sci. USA* **101**, 1945–1950 (2004).
- Low, K. H., Lim, C., Ko, H. W. & Edery, I. Natural variation in the splice site strength of a clock gene and species-specific thermal adaptation. *Neuron* **60**, 1054–1067 (2008).
- United States Naval Observatory (USNO) Astronomy Application Department. *Rise, Set, and Twilight Definitions* (http://aa.usno.navy.mil/faq/docs/RST_defs.php) (2011).
- Tauber, E. *et al.* Natural selection favors a newly derived *timeless* allele in *Drosophila melanogaster*. *Science* **316**, 1895–1898 (2007).
- Hamblen-Coyle, M. J., Wheeler, D. A., Rutla, J. E., Rosbash, M. & Hall, J. C. Behavior of period-altered rhythm mutants of *Drosophila* in light:dark cycles. *J. Insect Behav.* **5**, 417–446 (1992).
- Helfrich, C. & Engelmann, W. Evidences for circadian rhythmicity in the *per*⁰ mutant of *Drosophila melanogaster*. *Z. Naturforsch. C* **42**, 1335–1338 (1987).
- Dowse, H. B. & Ringo, J. M. Further evidence that the circadian clock in *Drosophila* is a population of coupled ultradian oscillators. *J. Biol. Rhythms* **2**, 65–76 (1987).
- Yoshii, T., Fujii, K. & Tomioka, K. Induction of *Drosophila* behavioral and molecular circadian rhythms by temperature steps in constant light. *J. Biol. Rhythms* **22**, 103–114 (2007).
- Currie, J., Goda, T. & Wijnen, H. Selective entrainment of the *Drosophila* circadian clock to daily gradients in environmental temperature. *BMC Biol.* **7**, 49 (2009).
- Rieger, D. *et al.* The fruit fly *Drosophila melanogaster* favors dim light and times its activity peaks to early dawn and late dusk. *J. Biol. Rhythms* **22**, 387–399 (2007).
- Bachleitner, W., Kempinger, L., Wulbeck, C., Rieger, D. & Helfrich-Forster, C. Moonlight shifts the endogenous clock of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **104**, 3538–3543 (2007).
- Kempinger, L., Dittmann, R., Rieger, D. & Helfrich-Forster, C. The nocturnal activity of fruit flies exposed to artificial moonlight is partly caused by direct light effects on the activity level that bypass the endogenous clock. *Chronobiol. Int.* **26**, 151–166 (2009).
- Netabach, M. N. *et al.* Electrical hyperexcitation of lateral ventral pacemaker neurons desynchronises downstream circadian oscillators in the fly circadian circuit and induces multiple behavioural period. *J. Neurosci.* **26**, 479–489 (2006).
- Miyasako, Y., Umezaki, Y. & Tomioka, K. Separate sets of cerebral clock neurons are responsible for light and temperature entrainment of *Drosophila* circadian locomotor rhythms. *J. Biol. Rhythms* **22**, 115–126 (2007).
- Yoshii, T., Vanin, S., Costa, R. & Helfrich-Forster, C. Synergic entrainment of *Drosophila*'s circadian clock by light and temperature. *J. Biol. Rhythms* **24**, 452–464 (2009).
- Zhang, Y., Liu, Y., Bilodeau-Wentworth, D., Hardin, P. E. & Emery, P. Light and temperature control the contribution of specific DN1 neurons to *Drosophila* circadian behavior. *Curr. Biol.* **20**, 600–605 (2010).
- Zhang, L. *et al.* DN1(p) circadian neurons coordinate acute light and PDF inputs to produce robust daily behavior in *Drosophila*. *Curr. Biol.* **20**, 591–599 (2010).
- Busza, A., Murad, A. & Emery, P. Interactions between circadian neurons control temperature synchronization of *Drosophila* behavior. *J. Neurosci.* **27**, 10722–10733 (2007).
- Daan, S. *et al.* Lab mice in the field: unorthodox daily activity and effects of a dysfunctional circadian clock allele. *J. Biol. Rhythms* **26**, 118–129 (2011).
- Gatterman, R. *et al.* Golden hamsters are nocturnal in captivity but diurnal in nature. *Biol. Lett.* **4**, 253–255 (2008).
- Tanoue, S., Krishnan, P., Krishnan, B., Dryer, S. E. & Hardin, P. E. Circadian clocks in antennal neurons are necessary and sufficient for olfaction rhythms in *Drosophila*. *Curr. Biol.* **14**, 638–649 (2004).
- Rosato, E. & Kyriacou, C. P. Analysis of locomotor activity rhythms in *Drosophila*. *Nature Protocols* **1**, 559–568 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Bastianello for building the natural light simulator, and C. Helfrich-Forster for comments on the manuscript. This work was funded by grants from the European Community (the 6th Framework Project EUCLOCK no. 018741; to R.C. and C.P.K.), the Biotechnology and Biological Sciences Research Council and National Environmental Research Council (to C.P.K.), the Royal Society Wolfson Research Merit Award (to C.P.K.) and a Medical Research Council studentship (to E.W.G.). R.C. also thanks the Italian Space Agency (DCMC grant) and the Ministero dell'Università e delle Ricerche (MIUR).

Author Contributions C.P.K. and R.C. conceived the study. C.P.K. supervised the UK work, and R.C. the Italian component. S.B. conducted all the field experiments in the UK, and S.V., F.S. and M.P. the fieldwork in Italy. S.V. performed the laboratory simulations. S.M. performed the *per* splicing work, and P.M. the immunocytochemistry experiments. E.W.G. developed the software to analyse the behavioural data. Statistical analyses were performed by S.B., S.V. and C.P.K. C.P.K., S.V., S.B. and R.C. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.C. (rodolfo.costa@unipd.it).

METHODS

Fly strains. Strains WTALA and HU are natural isolates collected as isofemale lines from northern Italy (Alto Adige 46° 30' N) in 2004 and from Houton, The Netherlands (52° 2' N) in 2004, respectively. Canton-S is a standard laboratory strain. *per*⁰¹, *per*^S and *per*^L are congenic and were backcrossed repeatedly to a Cantonized *per* deficiency *Df(1)64j4* background by C.P.K. in the early 1980s. The genetic backgrounds of the other fly strains used are varied, and for most they are unknown.

In the laboratory, even at a constant 10 °C, we observed that *UAS-CycA* (ref. 29) was effectively driven by the *timgal4* driver and gave an arrhythmic locomotor profile compared to the control strain (Supplementary Fig. 10). All flies were raised to adulthood at 25 °C in the laboratory before being placed outside at 2–3 days of age. We confirmed that *UAS-hidUAS-rpr* was effective in eliminating the clock neurons under *tim*- and *Pdf-gal4* control using immunocytochemistry (data not shown).

Locomotor activity. Locomotor activity was monitored in suburban gardens well away from street lighting and from any major roads. Flies were placed in Trikinetics activity monitors, and shielded from direct sunlight and from rain (Supplementary Fig. 1). In Leicester this was accomplished by placing the monitors on a table within a child's plastic playhouse (Toys R Us) with approximate internal dimensions of 4 × 4 × 4 ft³, with open windows in a north–south direction that was shaded by an adjacent large tree. Bricks placed on the roof ensured that the house did not blow away during high winds. Cables from monitors were fed from the back of the playhouse into a computer that was placed inside an immediately adjacent extension to C.P.K.'s house. In Italy, monitors were placed under a penthouse roof in a garden roughly 4 km from the city centre of Treviso, again with no direct sunlight, with cables fed to a computer within the penthouse. Flies were monitored from March to November (Treviso) or from May to September (Leicester). Simultaneous and continuous monitoring of temperature, light intensity and humidity were made with recorders placed adjacent to the activity tubes. Circadian rhythmicity was assessed by spectral analysis and autocorrelation as reported previously³⁰.

Morning (*M*_{onset}) and afternoon or evening (*A*_{onset}, *E*_{onset}) onsets were determined for each fly on each day of every experiment (see Supplementary Fig. 11 for examples). *M*_{onset} was considered to be present when a bout of activity occurred after a period of rest during the dark phase, if continuous movement with a steady increase in activity (and no more than one half-hour time bin without any or lower activity than the previous bin) led to a peak, followed by a steady decrease in activity defining the offset. Activity bouts in the middle of the night (not leading continuously into the morning) were not considered. In general, the window of time that contained the morning bout was 3 h before or after the 30-min time bin during which the light intensity reached 1 lx. If there was no bout of activity consistent with the parameters described above, it was noted that there was no morning component for that fly on that day.

*E*_{onset} was considered to be present when a bout of activity occurred after a period of rest during the day. If a burst of activity occurred in the middle of the afternoon on warmer days, with another clear bout after it, the latter was considered to represent the *E*_{onset}, whereas the former represented the afternoon component. The evening and afternoon bouts of activity had to be composed of continuous movement with no more than one zero activity bin interspersed within, and with a steady increase and decrease in activity levels defining the onset/offset. If there was no bout of activity consistent with the parameters described above, it was noted that there was no evening peak for that fly on that day.

Moonlight measurements. Measurements of moonlight were made in Treviso and Leicester, using a LI-210SA Photometric Sensor (LI-COR) connected to a LI 1400 data logger. Five-minute readings were made four times during selected nights of full moon or new moon, and the minimum and maximum light levels were recorded, both in fully exposed full moonlight conditions and in the more sheltered positions in which the Trikinetics monitors were normally maintained. Average maximum nightly full moonlight levels were 0.23 ± 0.15 lx in fully exposed positions, in comparison with 0.07 ± 0.01 lx in sheltered conditions in Leicester (30 November 2009). The corresponding values in Treviso were 0.33 ± 0.07 lx in exposed positions and 0.03 ± 0.01 lx in sheltered conditions. At new moon, in both Leicester and Treviso, the light values were 0 lx.

Twilights. Twilights were obtained from the online database of the United States Naval Observatory (USNO) Astronomy Application Department³¹.

Natural temperature profiles. These were reproduced in the laboratory by using the Memmert ICP 700 incubator with a refrigeration unit. The Celsius 2007 computer program, running under the Windows XP operating system, was used for programming, controlling and documenting the Memmert ICP 700 using the USB interface. The incubator can generate smooth cycling temperature ranges without step-ups.

Natural light profiles. These were simulated by using a programmable daylight simulator based on a battery of white-light LEDs of different intensities that could be programmed to generate smooth cycling profiles (without step-ups) of light ranging from 0.005 lx (astronomical twilight) to 1,500 lx. The front-end programming allowed the simulation of any observed natural cycle of light intensity. The equipment was designed and built by EURITMI, a spin-out from the Electronic Engineering Unit at the Venetian Institute of Molecular Medicine (Padova, Italy).

Immunocytochemistry. Flies were collected every 3 h for a 24-h time course in natural conditions and fixed for 3 h in 4% paraformaldehyde. After three washes (15 min each) in PBS, the brains were dissected under a stereomicroscope (Leica MB6). Brains were washed five times for 6 min each in 0.3% PBST (PBS + 0.3% Triton X-100). They were then permeabilized for 10 min in 1% PBST and blocked for 2 h with 1% BSA. After blocking, the brains were incubated for 3 days at 4 °C, in the primary antibody diluted in 0.1% BSA in 0.3% PBST. They were then washed five times for 6 min each in 0.3% PBST. After washing, they were blocked again for 1 h with 1% BSA. The brains were then incubated overnight at 4 °C in the appropriate secondary antibody. The antibodies used for the immunocytochemistry experiments were anti-PDF (Developmental Studies Hybridoma Bank, dilution of 1:5,000), anti-PER (from R. Stanewsky; 1:2,500), and anti-mouse and anti-rabbit (both from Invitrogen; 1:500).

Brains were mounted onto slides (VWR) with a drop of mounting medium (Vectashield; Vecta Laboratories, Inc.) and covered with a coverslip 0.1 mm thick (VWR). The slides were observed under a Nikon 80i microscope with a QICAM Fast Camera. Individual images were taken of planes at different depths to create a *z*-series for each brain. The size of the sections forming a *z*-series was 1.0 ± 0.2 µm. The images were viewed and quantified with ImageJ version 1.42g (<http://rsb.info.nih.gov/ij/>). The average pixel intensity for each neuron was measured together with the signal from its corresponding background area. The final amount of signal was calculated using the formula $\text{intensity} = 100 \times (\text{signal} - \text{background}) / \text{background}$. The values of intensity obtained were normalized using the higher value as 100% and then plotted with Microsoft Excel 2003 or OriginPro 8.0, and statistical analyses were performed with Statistica 8 (StatSoft).

RNA quantification. Flies were collected every 3 h for a 24-h time course in natural conditions and fixed in liquid nitrogen. Heads were removed and total RNA was recovered and extracted using Trizol Reagent (Gibco) as recommended by the kit protocol. Reverse transcription was initiated on total RNA, after any contaminating genomic DNA was removed with DNase (Promega), using the SuperscriptII (Invitrogen) reverse transcriptase with the 17-bases oligo(dT). The reaction was performed for 1 h at 42 °C and for 15 min at 75 °C. *per*^{spliced}, *per*^{unspliced} forms and the control gene *Cbp20* (*Cap binding protein 20*) were amplified by PCR with a PTC-100 Peltier Thermal Cycler (MJ Research) using primers listed below. PCR was performed as follows: initial denaturation at 95 °C for 3 min, then 28 cycles consisting of 95 °C for 1 min, 62.1 °C for 1 min, and 72 °C for 45 s. The reaction was completed by an elongation step of 10 min at 72 °C. Amplifications were carried out in 20-µl reaction mixtures containing 75 ng of cDNA target, 1 µl of each primer (10 µM), 1.6 µl of dNTPs (2 mM), 4 µl of Green Buffer (5×) and 0.4 µl of GoTaq DNA polymerase (Promega). Primers used were as follows: *per*, 5'-AAGACGAGCCGGCTCCAG-3' (forward; base pairs 6421–6440; NCBI X03636) and 5'-TCTACATTATCCTCGGCTTGC-3' (reverse; base pairs 7201–7221; NCBI X03636); *cbp20*, 5'-GTCTGATTCGTGTGGACTGG-3' (forward; base pairs 540–559; FlyBase ID FBgn0022943) and 5'-CAACAGTTTGCCATAACCC-3' (reverse; base pairs 653–672; FlyBase ID FBgn0022943).

PCR products were revealed on a 2% agarose (Eurobio) gel under ultraviolet radiation. Images were collected with the Quantity One 4.6 (Bio-Rad). The final quantification was obtained with ImageJ software by applying the formula $\text{intensity}_{\text{RNA(pixel)}} = \text{band intensity of } per \text{ spliced} / per \text{ unspliced} / \text{band intensity of } cbp20$. Each band's intensity was corrected for background noise. For each collection, three independent RNA extractions, reverse transcriptions, amplifications and quantifications were performed. The results were normalized and analysed with Statistica 8 (StatSoft). The plots were obtained with OriginPro 8.0 software.

Patterns and rates of exonic *de novo* mutations in autism spectrum disorders

Benjamin M. Neale^{1,2}, Yan Kou^{3,4}, Li Liu⁵, Avi Ma'ayan³, Kaitlin E. Samocha^{1,2}, Aniko Sabo⁶, Chiao-Feng Lin⁷, Christine Stevens², Li-San Wang⁷, Vladimir Makarov^{4,8}, Paz Polak^{2,9}, Seungtae Yoon^{4,8}, Jared Maguire², Emily L. Crawford¹⁰, Nicholas G. Campbell¹⁰, Evan T. Geller⁷, Otto Valladares⁷, Chad Schafer⁵, Han Liu¹¹, Tuo Zhao¹¹, Guiqing Cai^{4,8}, Jayon Lihm^{4,8}, Ruth Dannenfelser³, Omar Jabado¹², Zuleyma Peralta¹², Uma Nagaswamy⁶, Donna Muzny⁶, Jeffrey G. Reid⁶, Irene Newsham⁶, Yuanqing Wu⁶, Lora Lewis⁶, Yi Han⁶, Benjamin F. Voight^{2,13}, Elaine Lim^{1,2}, Elizabeth Rossin^{1,2}, Andrew Kirby^{1,2}, Jason Flannick³, Menachem Fromer^{1,2}, Khalid Shakir², Tim Fennell², Kiran Garimella², Eric Banks², Ryan Poplin², Stacey Gabriel², Mark DePristo², Jack R. Wimbish¹⁴, Braden E. Boone¹⁴, Shawn E. Levy¹⁴, Catalina Betancur¹⁵, Shamil Sunyaev^{2,9}, Eric Boerwinkle^{6,16}, Joseph D. Buxbaum^{4,8,12,17}, Edwin H. Cook Jr¹⁸, Bernie Devlin¹⁹, Richard A. Gibbs⁶, Kathryn Roeder⁵, Gerard D. Schellenberg⁷, James S. Sutcliffe¹⁰ & Mark J. Daly^{1,2}

Autism spectrum disorders (ASD) are believed to have genetic and environmental origins, yet in only a modest fraction of individuals can specific causes be identified^{1,2}. To identify further genetic risk factors, here we assess the role of *de novo* mutations in ASD by sequencing the exomes of ASD cases and their parents ($n = 175$ trios). Fewer than half of the cases (46.3%) carry a missense or nonsense *de novo* variant, and the overall rate of mutation is only modestly higher than the expected rate. In contrast, the proteins encoded by genes that harboured *de novo* missense or nonsense mutations showed a higher degree of connectivity among themselves and to previous ASD genes³ as indexed by protein-protein interaction screens. The small increase in the rate of *de novo* events, when taken together with the protein interaction results, are consistent with an important but limited role for *de novo* point mutations in ASD, similar to that documented for *de novo* copy number variants. Genetic models incorporating these data indicate that most of the observed *de novo* events are unconnected to ASD; those that do confer risk are distributed across many genes and are incompletely penetrant (that is, not necessarily sufficient for disease). Our results support polygenic models in which spontaneous coding mutations in any of a large number of genes increases risk by 5- to 20-fold. Despite the challenge posed by such models, results from *de novo* events and a large parallel case-control study provide strong evidence in favour of *CHD8* and *KATNAL2* as genuine autism risk factors.

In spite of the substantial heritability, few genetic risk factors for ASD have been identified^{1,2}. Copy number variants (CNVs), in particular *de novo* and large events spanning multiple genes, have been identified as conferring risk^{4,5}. Although these CNVs provide important leads to underlying biology, they rarely implicate single genes, are rarely fully penetrant, and many confer risk to a broad range of conditions including intellectual disability, epilepsy and schizophrenia⁶. There are also documented instances of rare single nucleotide variants (SNVs) that are highly penetrant for ASD³.

Large-scale genetic studies make clear that the origins of ASD risk are multifarious, and recent estimates based on CNV data put the

number of independent risk loci in the hundreds⁵. Yet knowledge regarding specific risk-determining genes and the overall genetic architecture for ASD remains incomplete. Although new sequencing technologies provide a catalogue of most variation in the genome, the profound locus heterogeneity of ASD makes it challenging to distinguish variants that confer risk from the background noise of inconsequential SNVs. *De novo* variation, being less frequent and potentially more deleterious, could offer insights into risk-determining genes. Accordingly, we sought to evaluate carefully the observed rate and consequence of *de novo* point mutations in the exomes of ASD subjects.

We performed exome sequencing of 175 ASD probands and their parents across five centres with multiple protocols and validation techniques (Supplementary Information). We used a sensitive and specific analytical pipeline based on current best practices⁷⁻⁹ to analyse all data and observed no heterogeneity of mutation rate across centres.

In the entire sample, we observed 161 coding region point mutations (101 missense, 50 silent and 10 nonsense), with an additional two conserved splice site (CSS) SNVs and six frameshift insertions/deletions (indels) validated and included in pathway analyses (Supplementary Table 1).

To determine whether the rate of coding region point mutations was elevated, we estimated the mutation rate in light of coverage and base context using two parallel approaches (Supplementary Information). On the basis of both models, the exome target should have a significantly increased ($\sim 30\%$) mutation rate compared to the genome. Conservatively, by assuming the low end of the estimated mutation rate from recent whole-genome data (1.2×10^{-8})¹⁰, we estimate a mutation rate of 1.5×10^{-8} for the exome sequence captured here. The observed point mutation rate of 0.92 per exome is slightly but not significantly elevated versus expectation (Table 1) and is insensitive to adjustment for lower coverage regions (Supplementary Information). Indeed our rate is similar to that of ref. 11.

Per-family events were distributed exquisitely according to the Poisson distribution (Table 1), suggesting limited variation in the underlying rate of *de novo* mutation in ASD families. The relative rates

¹Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA. ²Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. ³Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁴Seaver Autism Center for Research and Treatment, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁵Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15232, USA. ⁶Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA. ⁷Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. ⁸Department of Psychiatry, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁹Division of Genetics, Department of Medicine Brigham & Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. ¹⁰Vanderbilt Brain Institute, Departments of Molecular Physiology & Biophysics and Psychiatry, Vanderbilt University, Nashville, Tennessee 37232, USA. ¹¹Biostatistics Department and Computer Science Department, Johns Hopkins University, Baltimore, Maryland 21205, USA. ¹²Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York 10029, USA. ¹³Department of Pharmacology, University of Pennsylvania, Perelman School of Medicine, Philadelphia, Pennsylvania 19104, USA. ¹⁴HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA. ¹⁵INSERM U952 and CNRS UMR 7224 and UPMC Univ Paris 06, 75005 Paris, France. ¹⁶Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas 77030, USA. ¹⁷Friedman Brain Institute, Mount Sinai School of Medicine, New York, New York 10029, USA.

¹⁸Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois 60608, USA. ¹⁹Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15213, USA.

Table 1 | Distribution of events per family

| Events per family | All ASD trios | | Random mut. exp.‡ |
|-------------------|---------------|-------|-------------------|
| | Exon DN SNVs* | Exp.† | |
| 0 | 71 | 69.7 | 73.2 |
| 1 | 62 | 64.2 | 63.8 |
| 2 | 28 | 29.5 | 27.8 |
| 3 | 10 | 9.1 | 8.1 |
| 4 | 2 | 2.1 | 1.8 |
| 5 | 1 | 0.4 | 0.3 |
| Mean | | 0.920 | 0.871 |

* Exon DN SNVs include all single nucleotide variants in coding sequence but excludes indels and intronic variants.

† The expected distribution of number of trios with a given event count as determined by the Poisson.

‡ Random mut. exp. is the expectation for 175 trios based on the sequence-context mutation rate model M1 (Supplementary Information) based on the count of the number of trios that have at least 10 × coverage.

of ‘functional’ (missense, nonsense, CSS and read-through) versus silent changes did not deviate from expectation (Table 2). We did, however, observe ten nonsense mutations (6.2%), which exceeded expectation (3.3%) (one-tailed $P = 0.04$; Supplementary Information).

We examined missense mutations using PolyPhen-2 scores¹² to measure severity, as some missense variants can severely affect function¹³. These scores showed no deviation from random expectation. The observed PolyPhen-2 scores clearly deviate from standing variation in the parents (Table 2), but such variation, even the rarest category, has survived selective pressure and so is inappropriate for comparison to *de novo* events.

We observed three genes with two *de novo* mutations: *BRCA2* (two missense), *FAT1* (two missense) and *KCNMA1* (one missense, one silent). A gene with two or more non-synonymous *de novo* hits across a panel of trios might indicate strong candidacy. However, simulations (Supplementary Information) show that two such hits are inadequate to define a gene as a conclusive risk factor given the number of observed events in the study.

From analyses of secondary phenotypes (Supplementary Tables 2 and 3), the most striking result is that paternal and maternal age, themselves highly correlated ($r^2 = 0.679$, P -value < 0.0001), each strongly predicts the number of *de novo* events per offspring (paternal age, $P = 0.0013$; maternal age, $P = 0.000365$), consistent with aggregating mutations in germ cells in the paternal line¹⁴. Consistent with a liability threshold model, there is an increased rate of *de novo* mutation in female versus male cases (1.214 for females versus 0.914 for males); however, the difference is not significant, owing to limited sample size. Considering phenotypic correlates, we observed no rate difference between subjects with strict autism versus those with a broader ASD classification, between positive and negative family history, or any significant effect of *de novo* mutation on verbal, non-verbal or full-scale IQ (Supplementary Table 3).

Given that hundreds of loci are apparently involved in autism⁵ and *de novo* mutations therein affect ASD risk, we modelled different numbers of risk genes and penetrances (Supplementary Information) and show that a model of hundreds of genes with high penetrance mutations is excluded by our data; however, more modest contributions of *de novo* variants are not. For example, up to 20% of cases

Table 2 | Rates of mutation annotation given variant type

| Type of <i>de novo</i> mutation | <i>De novo</i> (%)* | Random <i>de novo</i> (%) | Singletons (%)† | Doubletons (%)† | ≥3 (%)† |
|---|---------------------|---------------------------|-----------------|-----------------|---------|
| Missense | 62.7 | 66.1 | 59.5 | 55.4 | 48.8 |
| Nonsense | 6.2 | 3.3 | 1.2 | 0.8 | 0.4 |
| Synonymous | 31.1 | 30.6 | 39.3 | 43.8 | 50.8 |
| PolyPhen-2 missense classification | | | | | |
| Benign | 35.0 | 35.9 | 46.6 | 51.3 | 63.4 |
| Possibly damaging | 21.0 | 18.9 | 18.8 | 17.7 | 15.1 |
| Probably damaging | 44.0 | 45.2 | 34.7 | 31.0 | 21.4 |

* All indels and failing variants were removed.

† Singletons, doubletons and ≥3 (copies) are only those variants called in 192 parents.

carrying a *de novo* event conferring a 10- or 20-fold increased risk is consistent with these data (Supplementary Table 4). Thus, our data are consistent with either chance mutation or a modest role for *de novo* mutations on risk. Importantly, a single deleterious event is unlikely to fully explain disease in a patient.

We therefore posed two questions of the group of genes harbouring *de novo* functional mutations: do the protein products of these genes interact with each other more than expected, and are they unusually enriched in, or connected to, previous curated lists of ASD-implicated genes? Using an *in silico* approach (DAPPLE)¹⁵, the protein–protein connectivity defined by InWeb¹⁶ in the set of 113 genes harbouring functional *de novo* mutations was evaluated. These analyses (Fig. 1) showed significantly greater connectivity among the *de novo* identified proteins than would be expected by chance ($P < 0.001$) (Supplementary Information).

Querying previously defined, manually curated lists of genes³ associated with high risk for ASD with or without intellectual disability (Supplementary Table 5), and high-risk intellectual disability genes (Supplementary Table 6), we asked whether there was significant enrichment for *de novo* mutations in these genes. Five genes with functional *de novo* events were previously associated with ASD and/or intellectual disability (*STXBP1*, *MEF2C*, *KIRREL3*, *RELN* and *TUBA1A*); for four of these genes (all but *RELN*) the previous evidence indicated autosomal dominant inheritance.

We then assessed the average distance (D_i , Supplementary Fig. 2) of the *de novo* coding variants in brain-expressed genes (see supplement) to the ASD/intellectual disability list using a protein–protein interaction background network. To enhance power, data from a companion study¹¹ were used, including the observed silent *de novo* variants and *de novo* variants in unaffected siblings as comparators. The average distance for non-synonymous variants was significantly smaller for the case set than the comparator set (3.66 ± 0.42 versus 3.78 ± 0.59 ; permutation $P = 0.033$) (Supplementary Fig. 3). Much of this signal comes from 31 synaptic genes identified by three large-scale synaptic proteomic studies ($D_i = 3.47 \pm 0.46$ versus 3.57 ± 0.60 ; permutation $P = 0.084$) (Fig. 2; see also Supplementary Fig. 4 for the complete data). Taken in total, these independent gene set analyses, along with the modest enrichment of *de novo* variants over background rates in

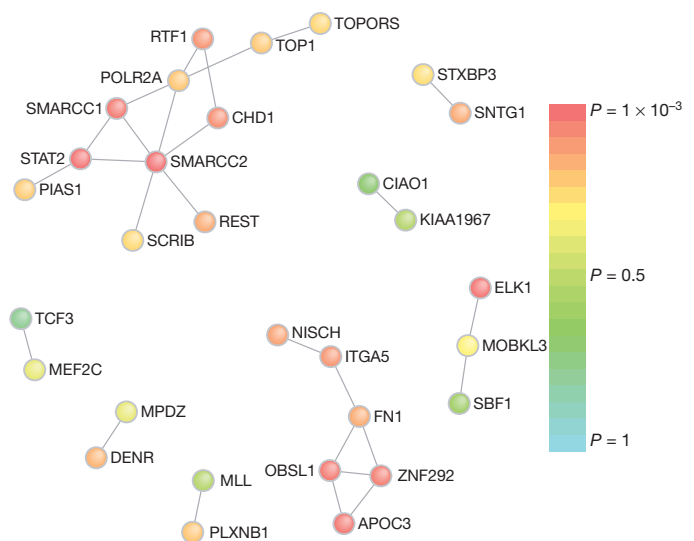


Figure 1 | Protein–protein interaction for genes with an observed functional *de novo* event. Direct protein connections from InWeb, restricting to genes harbouring *de novo* mutations for DAPPLE analysis. Two extensive networks are identified: the first is centred on SMARCC2 with 12 connections across 11 genes; the second is centred on FN1 with 7 connections across 6 genes. The P value for each gene having as many connections as those observed is indicated by node colour.

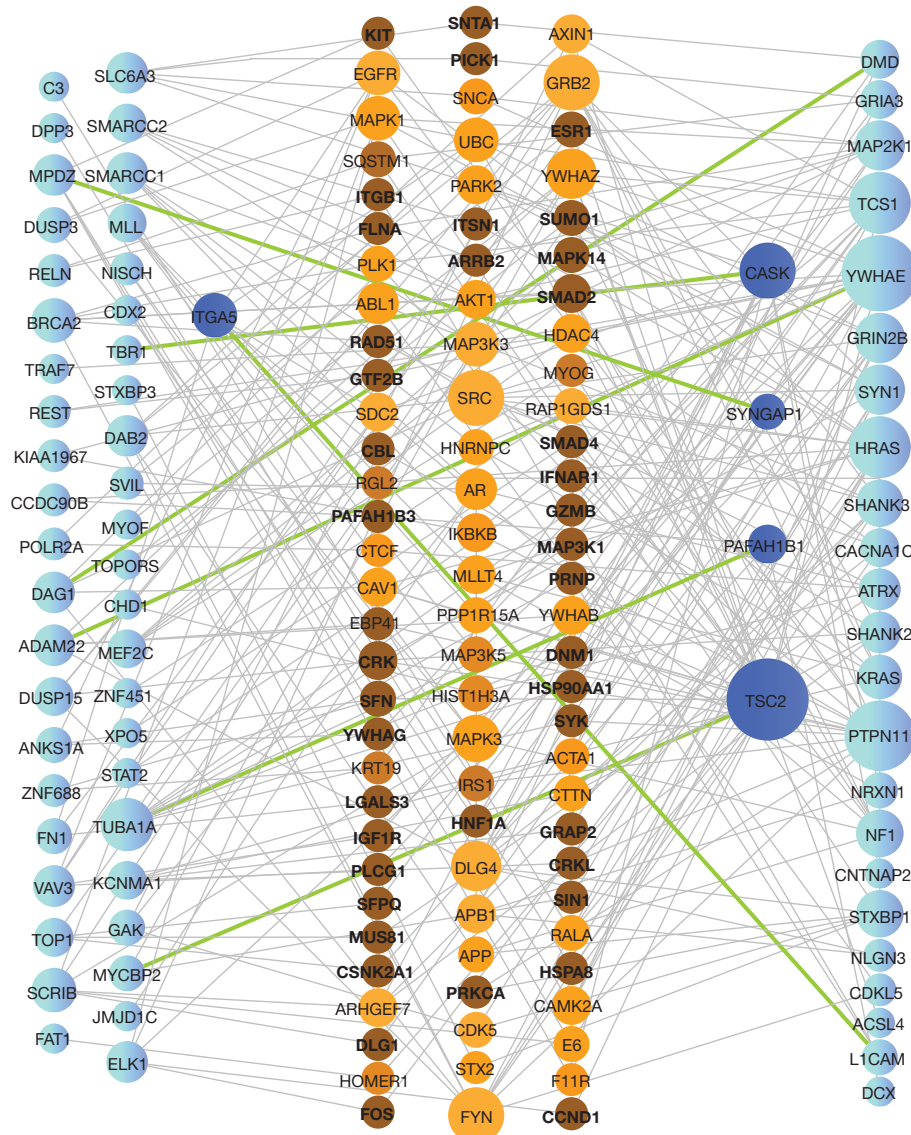


Figure 2 | Direct and indirect protein–protein interaction for genes with a functional *de novo* event and previous ASD genes. PPI network analysis for *de novo* variants and 31 previous synaptic ASD genes (see Supplementary Information). Nodes are sized based on connectivity. Genes harbouring *de novo* variants (left) and previous ASD genes (right) are coloured blue, with dark blue blue nodes representing genes that belong to one of these lists and are also

intermediate proteins. Intermediate proteins (centre) are coloured in shades of orange based on a *P* value computed using a proportion test, where a darker colour represents a lower *P* value. Green edges represent direct connections between genes harbouring *de novo* variants (left) and previous ASD genes. All other edges, connecting to intermediate proteins, are shown in grey.

ASD, indicate that a proportion of the *de novo* events observed in this study probably contribute to autism risk.

Using whole-exome sequencing of autism trios, we demonstrate a rate, functional distribution and predicted impact of *de novo* mutation largely consistent with chance mutational processes governed by sequence context. This lack of significant deviation from random mutational processes indicates a more limited role for the contribution of *de novo* mutations to ASD pathogenesis than has previously been suggested¹⁷, and specifically highlights the fact that observing a single *de novo* mutation, even an apparently ‘severe’ loss-of-function allele, is insufficient to implicate a gene as a risk factor. Yet the pathway analyses presented here assert that the overall set of genes hit with functional *de novo* mutations is not random and that these genes are biologically related to each other and to previously identified ASD/intellectual disability candidate genes. Modelling the *de novo* mutational process under a range of genetic models reveals that some models are inconsistent with the observed data—for example, 100 rare, fully penetrant Mendelian genes similar to Rett’s syndrome—whereas

others are not inconsistent, such as spontaneous ‘functional’ mutation in hundreds of genes that would increase risk by 10- or 20-fold (Supplementary Table 4). Models that fit the data are consistent with the relative risks estimated for most *de novo* CNVs⁵ and suggest that *de novo* SNVs, like most CNVs, often combine with other risk factors rather than fully cause disease. Furthermore, these models indicate that *de novo* SNV events will probably explain <5% of the overall variance in autism risk (Supplementary Table 4).

Considering the two companion papers^{11,18}, 18 genes with two functional *de novo* mutations are observed in the complete data. Using simulations, 11.91 genes on average harbour functional mutations by chance (Supplementary Table 7). Thus, a set of 18 genes with two or more hits is not quite significant (*P* = 0.063). Matching loss-of-function variants, however, at *SCN2A*, *KATNAL2* and *CHD8* (Supplementary Table 7) are unlikely to occur by chance because of the expected very low rate of *de novo* nonsense, splice and frameshift variants. We evaluated these strong candidates further using exome sequencing on 935 cases and 870 controls, and at both *KATNAL2* and

CHD8 three additional loss-of-function mutations were observed in cases with none in controls. No additional loss-of-function mutations were seen at *SCN2A* in the case-control data, but a new splice site *de novo* event has been validated in an additional autism case while this paper was in press, strengthening the evidence for this gene as relevant to autism. Using data from more than 5,000 individuals in the NHLBI Exome Variant Server (<http://evs.gs.washington.edu/EVS/>) as additional controls, three loss-of-function mutations were seen in *KATNAL2* but none in *CHD8*, making the additional observation of three *CHD8* loss-of-function mutations in our cases significant evidence ($P < 0.01$) of this being a genuine autism susceptibility gene. Not all genes with double hits are nearly so promising (Supplementary Information and Supplementary Tables 8 and 9), supporting the estimate above that most of such observations are simply chance events. Overall, these data underscore the challenge of establishing individual genes as conclusive risk factors for ASD, a challenge that will require larger sample sizes and deeper analytical integration with inherited variation.

METHODS SUMMARY

We ascertained probands using the Autism Diagnostic Interview-Revised (ADI-R), the Autism Diagnostic Observation Schedule-Generic (ADOS) and the DSM-IV diagnosis of a pervasive developmental disorder. All probands met criteria for autism on the ADI-R and either autism or ASD on the ADOS, except for the three subjects that were not assessed with the ADOS. All subjects provided informed consent and the research was approved by institutional human subjects boards.

For 175 trios, we performed exome capture and sequencing using either the Agilent 38Mb SureSelect v2 ($n = 118$), the NimbleGen Seq Cap EZ SR v2 ($n = 51$), or NimbleGen VCRome 2.1 (Baylor $n = 6$). After capture, another round of LM-PCR was performed to increase the quantity of DNA available for sequencing. All libraries were sequenced using an IlluminaHiSeq2000.

All sequence data were processed with Picard (<http://picard.sourceforge.net/>), which recalibrates quality scores and local realignment at known indels⁸ and BWA⁷ for mapping reads to hg19. SNPs were called using GATK^{8,9} for all trios jointly. Putative *de novo* mutations were identified restricting to sites passing standard filters and both parents were homozygous for the reference sequence and the offspring was heterozygous, and each genotype call was made confidently (see Supplementary Information).

All putative *de novo* events were validated by sequencing the carrier and both parents using Sanger sequencing methods (71 trios) or by using Sequenom MALDI-TOF (104 trios). All events were annotated using RefSeq hg19.

We modelled a Poisson process consistent with the mutation model and observed data. We varied the fraction of genes that influence risk, the probability of a functional variant, and the penetrance of said events.

We performed association tests using SKAT¹⁹, a generalization of C-alpha²⁰. Our primary analyses treat case-control data generated at Baylor and Broad sequencing centres separately (23 genes \times 2 sites), but we also performed mega- and meta-analyses (23 genes \times 2 methods).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 September 2011; accepted 6 March 2012.

Published online 4 April 2012.

1. Lichtenstein, P., Carlstrom, E., Rastam, M., Gillberg, C. & Anckarsater, H. The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am. J. Psychiatry* **167**, 1357–1363 (2010).
2. Hallmayer, J. *et al.* Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiatry* **68**, 1095–1102 (2011).
3. Betancur, C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res.* **1380**, 42–77 (2011).
4. Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368–372 (2010).
5. Sanders, S. J. *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70**, 863–885 (2011).
6. Sebat, J., Levy, D. L. & McCarthy, S. E. Rare structural variants in schizophrenia: a one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet.* **25**, 528–535 (2009).

7. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
8. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genet.* **43**, 491–498 (2011).
9. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
10. Conrad, D. F. *et al.* Variation in genome-wide mutation rates within and between human families. *Nature Genet.* **43**, 712–714 (2011).
11. Sanders, S. J. *et al.* *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* <http://dx.doi.org/10.1038/nature10945> (this issue).
12. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nature Methods* **7**, 248–249 (2010).
13. Kryukov, G. V., Pennacchio, L. A. & Sunyaev, S. R. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am. J. Hum. Genet.* **80**, 727–739 (2007).
14. Crow, J. F. The origins, patterns and implications of human spontaneous mutation. *Nature Rev. Genet.* **1**, 40–47 (2000).
15. Rossin, E. J. *et al.* Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet.* **7**, e1001273 (2011).
16. Lage, K. *et al.* A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc. Natl Acad. Sci. USA* **105**, 20870–20875 (2008).
17. O’Roak, B. J. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nature Genet.* **43**, 585–589 (2011).
18. O’Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* <http://dx.doi.org/10.1038/nature10989> (this issue).
19. Wu, M. C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* **89**, 82–93 (2011).
20. Neale, B. M. *et al.* Testing for an unusual distribution of rare variants. *PLoS Genet.* **7**, e1001322 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was directly supported by NIH grants R01MH089208 (M.J.D.), R01 MH089025 (J.D.B.), R01 MH089004 (G.D.S.), R01MH089175 (R.A.G.) and R01 MH089482 (J.S.S.), and supported in part by NIH grants P50 HD055751 (E.H.C.), R01 MH057881 (B.D.) and R01 MH061009 (J.S.S.). Y.K., G.C. and S.Y. are Seaver Fellows, supported by the Seaver Foundation. We thank T. Lehner, A. Felsenfeld and P. Bender for their support and contribution to the project. We thank S. Sanders and M. State for discussions on the interpretation of *de novo* events. We thank D. Reich for comments on the abstract and message of the manuscript. We thank E. Lander and D. Altshuler for comments on the manuscript. We acknowledge the assistance of M. Potter, A. McGrew and G. Crockett without whom these studies would not be possible, and Center for Human Genetics Research resources: Computational Genomics Core, Genetic Studies Ascertainment Core and DNA Resources core, supported in part by NIH NCR grant UL1 RR024975, and the Vanderbilt Kennedy Center for Research on Human Development (P30 HD015052). This work was supported in part by R01MH084676 (S.S.). We acknowledge the clinicians and organizations that contributed to samples used in this study and the particular support of the Mount Sinai School of Medicine, University of Illinois-Chicago, Vanderbilt University, the Autism Genetics Resource Exchange and the institutions of the Boston Autism Consortium. We acknowledge A. Estes and G. Dawson for patient collection/characterization. We acknowledge partial support from U54 HG003273 (R.A.G.) and U54 HG003067 (E. Lander). J.D.B., B.D., M.J.D., R.A.G., A.S., G.D.S. and J.S.S. are lead investigators in the Autism Sequencing Consortium (ASC). The ASC is comprised of groups sharing massively parallel sequencing data in autism. Finally, we are grateful to the many families, without whose participation this project would not have been possible.

Author Contributions Laboratory work: A.S., C.St., G.C., O.J., Z.P., J.D.B., D.M., I.N., Y.W., L.L., Y.H., S.G., E.L.C., N.G.C. and E.T.G. Data processing: B.M.N., K.E.S., E.L., A.K., J.F., M.F., K.S., T.F., K.G., E.Ba., R.P., M.DeP., S.G., S.Y., V.M., J.L., J.D.B., A.S., C.St., U.N., J.G.R., J.R.W., B.E.B., S.E.L., C.F.L., L.S.W. and O.V. Statistical analysis: B.M.N., L.L., K.E.S., C.Sh., B.F.V., J.M., E.R., S.S., P.P., Y.K., A.M., R.D., C.-F.L., L.-S.W., H.L., T.Z., E.Bo., R.A.G., J.D.B., C.B., E.H.C., J.S.S., G.D.S., B.D., K.R. and M.J.D. Principal Investigators/study design: E.Bo., R.A.G., E.H.C., J.D.B., K.R., B.D., G.D.S., J.S.S. and M.J.D. Y.K., L.L., A.M., K.E.S., A.S. and C.-F.L. contributed equally to this work. E.Bo., J.D.B., E.H.C., B.D., R.A.G., K.R., G.D.S., J.S.S. and M.J.D. are lead investigators of the ARRA Autism Sequencing Collaboration.

Author Information Data included in this manuscript have been deposited at dbGaP under accession number phs000298.v1.p1 and is available for download at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000298.v1.p1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.J.D. (mjdy@atgu.mgh.harvard.edu), J.D.B. (joseph.buxbaum@mssm.edu) or K.R. (kathryn.roeder@gmail.com).

METHODS

Phenotype assessment. Affected probands were assessed by research-reliable research personnel using Autism Diagnostic Interview-Revised (ADI-R), and the Autism Diagnostic Observation Schedule-Generic (ADOS) and DSM-IV diagnosis of a pervasive developmental disorder was made by a clinician. All probands met criteria for autism on the ADI-R and either autism or ASD on the ADOS, except for the three subjects from AGRE that were not assessed with the ADOS. In all, 85% of probands were classified with autism on both the ADI-R and ADOS. All subjects provided informed consent and the research was approved by institutional human subjects boards.

Exome sequencing, variant identification and *de novo* detection. Exome capture and sequencing was performed at each site using similar methods. Exons were captured using the Agilent 38 Mb SureSelect v2 (University of Pennsylvania and Broad Institute $n = 118$), the NimbleGen Seq Cap EZ SR v2 (Mt Sinai School of Medicine, Vanderbilt University $n = 51$), or NimbleGen VCRome 2.1 (Baylor $n = 6$). After capture, another round of LM-PCR was performed to increase the quantity of DNA available for sequencing. All libraries were sequenced using an IlluminaHiSeq2000.

Sequence processing and variant calling was performed using a similar computational workflow at all sites. Data were processed with Picard (<http://picard.sourceforge.net/>), which uses base quality-score recalibration and local realignment at known indels⁸ and BWA⁷ for mapping reads to hg19. SNPs were called using GATK^{8,9} for all trios jointly. The variable sites that we have considered in analysis are restricted to those that pass GATK standard filters. From this set of variants, we identified putative *de novo* mutations as sites where both parents were homozygous for the reference sequence and the offspring was heterozygous and each genotype call was made confidently (see Supplementary Information).

Validation of *de novo* events. Putative *de novo* events were validated by sequencing the carrier and both parents using Sanger sequencing methods

(University of Pennsylvania, Mt Sinai School of Medicine, Vanderbilt University, Baylor Medical College) or by Sequenom MALDI-TOF genotyping of trios (Broad).

Gene annotation. All identified mutations were then annotated using RefSeq hg19. The functional impact of variants was assessed for all isoforms of each gene, with the most severe annotation taking priority. Splice site variants were identified as occurring within two base pairs of any intron/exon boundary.

Expectation of *de novo* mutation calculation. To calculate the expected *de novo* rate, we assessed the mutability of all possible trinucleotide contexts in the intergenic region of the human genome for variation in two fashions: fixed genomic differences compared to chimpanzee and baboon¹² and variation identified from the 1,000 Genomes project. The overall mutation rate for the exome was then determined by summing the probability of mutation for all bases in the exome that were captured successfully. We also determined the probability of each class functional mutation by summing the annotated variants.

Pathway analyses. We applied DAPPLE¹⁵, which uses the InWeb database¹⁶, to determine whether there is excess protein–protein interaction across the genes hit by a functional *de novo* event. We also assessed whether these genes were more closely connected to a list of ASD genes³.

Modelling *de novo* events. We modelled a Poisson process consistent with the expected distribution defined by the mutation model and with the observed data. We varied the fraction of genes that influence risk, the probability a variant in a gene would be functional, and the penetrance of functional *de novo* events. We also simulated a random set of *de novo* events to estimate the probability of hitting a gene multiple times.

Association analysis. We performed association tests using SKAT¹⁹, a generalization of C-alpha²⁰. Our primary analyses treat case–control data generated at Baylor and Broad sequencing centres separately (23 genes \times 2 sites), but we also performed mega- and meta-analyses (23 genes \times 2 methods).

Systematic discovery of structural elements governing stability of mammalian messenger RNAs

Hani Goodarzi^{1,2,†}, Hamed S. Najafabadi^{3,4,†}, Panos Oikonomou^{1,2,†}, Todd M. Greco², Lisa Fish⁵, Reza Salavati^{3,4,6}, Ileana M. Cristea² & Saeed Tavazoie^{1,2,†}

Decoding post-transcriptional regulatory programs in RNA is a critical step towards the larger goal of developing predictive dynamical models of cellular behaviour. Despite recent efforts^{1–3}, the vast landscape of RNA regulatory elements remains largely uncharacterized. A long-standing obstacle is the contribution of local RNA secondary structure to the definition of interaction partners in a variety of regulatory contexts, including—but not limited to—transcript stability³, alternative splicing⁴ and localization³. There are many documented instances where the presence of a structural regulatory element dictates alternative splicing patterns (for example, human cardiac troponin T) or affects other aspects of RNA biology⁵. Thus, a full characterization of post-transcriptional regulatory programs requires capturing information provided by both local secondary structures and the underlying sequence^{3,6}. Here we present a computational framework based on context-free grammars^{3,7} and mutual information² that systematically explores the immense space of small structural elements and reveals motifs that are significantly informative of genome-wide measurements of RNA behaviour. By applying this framework to genome-wide human mRNA stability data, we reveal eight highly significant elements with substantial structural information, for the strongest of which we show a major role in global mRNA regulation. Through biochemistry, mass spectrometry and *in vivo* binding studies, we identified human HNRPA2B1 (heterogeneous nuclear ribonucleoprotein A2/B1, also known as HNRNP A2B1) as the key regulator that binds this element and stabilizes a large number of its target genes. We created a global post-transcriptional regulatory map based on the identity of the discovered linear and structural *cis*-regulatory elements, their regulatory interactions and their target pathways. This approach could also be used to reveal the structural elements that modulate other aspects of RNA behaviour.

To isolate stability from other aspects of mRNA behaviour, we performed whole-genome mRNA stability measurements by incubating human MDA-MB-231 breast cancer cells in the presence of 4-thiouridine, which is efficiently incorporated into cellular RNA. Subsequently, 4-thiouridine-labelled transcripts were captured and quantified at different time-points after the removal of 4-thiouridine from the growth medium. We calculated a relative decay rate for each transcript based on the rate at which 4-thiouridine-labelled transcripts, in the absence of 4-thiouridine in the media, are replaced by newly synthesized unlabelled mRNAs in the population (Supplementary Fig. 1). These measurements were then used to identify the putative *cis*-regulatory elements (linear and structural) that underlie transcript stability. A number of methods have been previously introduced for discovering structural motifs mainly based on free energy minimization, local sequence alignments or a combination of both alignments and secondary structure predictions^{3,6,8}. However, the extent to which

these *in silico* predictions reflect stable *in vivo* molecular conformations has not been fully explored⁹. In fact, the RNA binding proteins and complexes that interact with their target transcripts may facilitate the formation of secondary structures *in vivo*. Thus, we sought to bypass the need for predicting thermodynamically stable secondary structures by efficiently enumerating a large space of potential structural motifs. We developed TEISER (Tool for Eliciting Informative Structural Elements in RNA), a framework for identifying the structural motifs that are informative of whole-genome measurements across all the transcripts. In this approach, structural motifs are defined in terms of context-free grammars⁷ (CFGs) that represent hairpin structures as well as primary sequence information (see Methods and Supplementary Fig. 2). TEISER employs mutual information to measure the regulatory consequences of the presence or absence of each of roughly 100 million different seed CFGs (see Methods). Mutual information is a robust non-parametric measure that reveals general dependencies across discrete or continuous measurements^{2,10}. For example, when applied to the transcript stability data, TEISER captures the dependency between the stability of each mRNA and the presence or absence of a given structural motif in its 5' and 3' untranslated regions (UTRs). TEISER, subsequently, uses these measurements to choose and further refine the most informative motifs, and performs a series of statistical tests—for example, randomization-based statistics and jackknifing tests—to achieve very low (<0.01) false-discovery rates (see Methods and Supplementary Fig. 2).

Application of TEISER to the mRNA stability measurements in MDA-MB-231 cells revealed eight strong structural motif predictions that passed our statistical tests aimed at finding the most likely elements causally involved in mRNA stability (Fig. 1 and Supplementary Fig. 3). Apart from being highly informative of mRNA stability measurements, these putative regulatory elements show a variety of other characteristics that support their functionality. For example, four of the discovered motifs are also informative of transcript stability measurements in mouse¹¹ (Supplementary Fig. 4a). Furthermore, these motifs are highly conserved between human and mouse genomes (see Methods and Supplementary Fig. 3) and are also informative of co-expression clusters discovered across independent whole-genome data sets (Supplementary Fig. 4b).

Among the putative structural motifs discovered by TEISER, we chose sRSM1 (structural RNA stability motif 1)—the most statistically significant 3' UTR element (*z*-score = 122)—for further analysis. In order to probe the functionality of sRSM1 instances across the genome, we performed *in vivo* titration experiments using synthetic oligonucleotides^{10,12}. Upon transfecting MDA-MB-231 cells with decoy RNA molecules harbouring sRSM1 instances (Supplementary Fig. 5), we observed a notable reduction in the level of endogenous transcripts that carried this motif, in comparison to their level in the control cells transfected with scrambled RNA molecules (Fig. 2). This global

¹Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08540, USA. ²Department of Molecular Biology, Princeton University, Princeton, New Jersey 08540, USA.

³Institute of Parasitology, McGill University, Montreal, Quebec H3G1Y6, Canada. ⁴McGill Centre for Bioinformatics, McGill University, Montreal, Quebec H3G1Y6, Canada. ⁵Laboratory of Systems Cancer Biology, Rockefeller University, New York, New York 10065, USA. ⁶Department of Biochemistry, McGill University, Montreal, Quebec H3G1Y6, Canada. [†]Present addresses: Department of Biochemistry and Molecular Biophysics, and Initiative in Systems Biology, Columbia University, New York, New York 10032, USA (H.G., P.O., S.T.); The Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada (H.S.N.).

In order to directly assess the regulatory consequences of modulating HNRPA2B1, we performed knock-down experiments followed by gene expression profiling. Consistent with our prior observations, HNRPA2B1 knock-down caused a significant decrease in the expression level of transcripts carrying sRSM1 (Fig. 3a). Stability measurements in the knock-down cells confirmed that the observed downregulation of these transcripts was in fact due to changes in stability (see Methods), with the transcripts carrying sRSM1 elements showing a marked increase in their corresponding relative decay rates (Fig. 3b).

In principle, our observations are consistent with a possible indirect role for HNRPA2B1—brought about, for instance, by a common partner that binds both HNRPA2B1 and sRSM1 sites. The direct interaction between HNRPA2B1 and its potential target genes can be tested through cross-linking and immunoprecipitation of HNRPA2B1,

which, through local ultraviolet photoreactivity of bases and amino acids, can detect direct physical interactions¹⁸. We expressed a tagged clone of HNRPA2B1 in MDA-MB-231 cells, and after ultraviolet-crosslinking, immunoprecipitated this protein and the target mRNA molecules that were bound to it. We then labelled the isolated RNA population and hybridized it to microarrays with the input total RNA as control (a method called RIP-chip¹⁹). We observed a highly significant enrichment of sRSM1 in the immunoprecipitated population (Fig. 3c). In order to reduce the background and better pinpoint the HNRPA2B1 binding sites, we treated the samples with nuclease before immunoprecipitation under denaturing conditions and sequenced the HNRPA2B1-bound RNA population (HITS-CLIP²⁰). We observed that sRSM1 elements were significantly enriched in the identified putative binding sites, in comparison with randomly selected sequences²¹ (Fig. 3d). These observations demonstrate that HNRPA2B1 directly

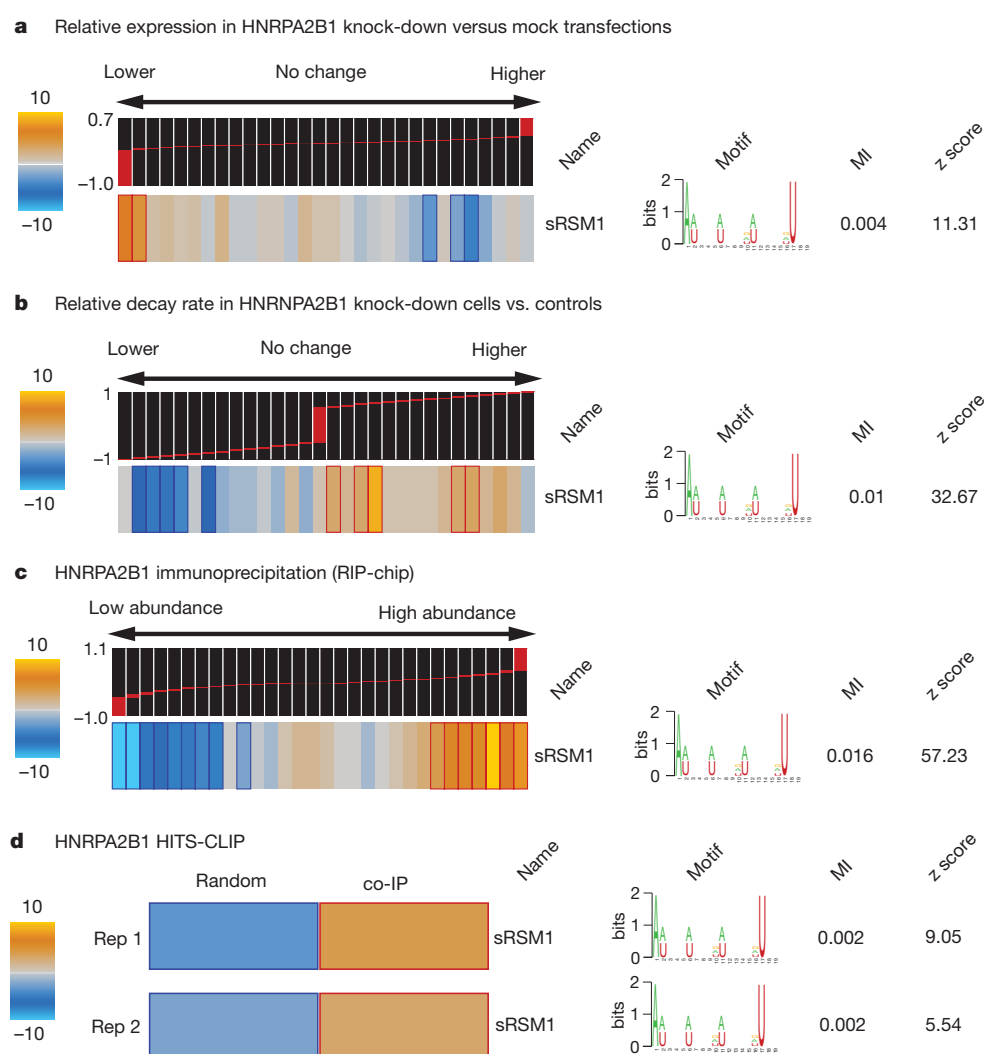


Figure 3 | HNRPA2B1 stabilizes transcripts through direct *in vivo* binding to sRSM1 structural motifs. **a**, Genome-wide expression levels were measured in HNRPA2B1 siRNA-transfected samples relative to mock-transfected controls. TEISER was used to capture the enrichment/depletion pattern of transcripts carrying sRSM1 across the relative expression values. Experiments were performed in triplicate, each with an independent siRNA targeting HNRPA2B1 and the resulting log ratios were averaged for each transcript. **b**, Transcript decay rates were compared in HNRPA2B1 knock-downs versus mock-transfected controls. These measurements were then analysed by TEISER to visualize the extent to which the decay rates of transcripts carrying sRSM1 elements were increased following HNRPA2B1 knock-down. **c**, Using ultraviolet-crosslinking followed by immunoprecipitation, mRNAs that bind

HNRPA2B1 were extracted and compared against the input mRNA population (RIP-chip). The log ratio calculated for each mRNA denotes its abundance in the immunoprecipitated sample relative to the input control. Bins to the right contain the mRNAs that were captured as interacting partners with HNRPA2B1. Similar to the prior examples, TEISER was used to show the enrichment/depletion pattern of transcripts carrying sRSM1 in their 3' UTRs. The values associated with each transcript were calculated as the average of log ratios from biological replicates. **d**, HNRPA2B1 binding sites were identified using immunoprecipitation followed by high-throughput sequencing (HITS-CLIP). Instances of the sRSM1 element are significantly enriched in these sites relative to a population of random sequences from 3' UTRs that are not represented in the sequenced population.

interacts with sRSM1 *in vivo* and acts to stabilize its target transcripts through this regulatory element. These transcripts, in turn, modulate a variety of cellular processes and pathways. For example, we observed a significant positive correlation between sRSM1 target transcripts and doubling-time in NCI-60 breast cancer cell lines (Fig. 4a). Indeed, knocking-down HNRPA2B1 resulted in a slight but significant increase in growth rate (by 10%, P -value $< 10^{-8}$), further highlighting the regulatory role of this global modulator in a key cellular process (Fig. 4b).

Revealing the detailed post-transcriptional regulatory code relies on the discovery of all the *cis*-regulatory elements that contribute to changes in transcript abundance. In addition to the sRSMs identified through TEISER, we also discovered a large diverse set of IRSMs (linear RNA stability motifs), including six known microRNA recognition sites, that are informative of transcript stability measurements (Supplementary Fig. 9). These motifs were identified by FIRE² (Finding Informative Regulatory Elements), a framework for discovering informative linear motifs. Combining these two approaches provided us with an extensive set of putative regulatory elements that cover both structural and primary sequence components. The next step in deciphering the post-transcriptional regulatory program involves the identification of target pathways that are potentially modulated by each element. Using iPAGE¹⁰ (Pathway Analysis of Gene Expression), we showed that our discovered elements probably target a diverse array of cellular processes and pathways (Supplementary Fig. 10). For example, the sRSM1 structural element is significantly enriched in the 3' UTRs of the genes involved in 'Notch signalling', while avoiding the UTRs of other pathways such as 'nucleosome assembly' (Supplementary Fig. 11). These results demonstrate that while post-transcriptional regulatory mechanisms are poorly

characterized, they have potentially far-reaching impact on specific cellular processes.

Regulatory programs often employ combinatorial interactions between various *cis*-regulatory elements to modulate gene expression^{2,22}. We used mutual information to reveal such potential interactions in the post-transcriptional regulatory programs governing mRNA stability (Supplementary Figs 12 and 13). For example, sRSM1 showed significant interactions with a number of structural and linear motifs, including sRSM8 and sRSM3 (Supplementary Fig. 11). These observed interactions might reflect cross-talk, or insulation, between the underlying regulatory processes that act upstream of these elements. The full map of such interactions (Supplementary Figs 14 and 15) reveals a complex network of motif-pathway relationships that set the stage for molecular dissection and predictive modelling of post-transcriptional regulation from sequence.

Whereas we have studied mRNA stability under normal and static conditions in a single cell line, the full regulatory program that governs mRNA stability is likely to involve a much richer repertoire of *cis*-regulatory elements operating within a more complex regulatory network. Also, although we have focused on transcript stability, our framework is general in concept and can be employed to study regulatory programs governing other aspects of RNA biology. For example, the established role of local secondary structures in shaping the splicing code^{4,23} suggests alternative splicing as a prominent area for analysis using this framework. The large repertoire of publicly available whole-genome expression data sets similarly offers a rich resource for identifying the post-transcriptional regulatory modules that underlie steady-state measurements.

METHODS SUMMARY

TEISER relies on calculating mutual information (MI) values between whole-genome measurements and millions of predefined structural motifs. The statistically significant motifs are then optimized and elongated through a heuristic search algorithm. The mRNA stability measurements were performed using a previously published method⁴. The decoy/scrambled experiments and siRNA knock-downs were performed using lipofectamin 2000 reagent (Invitrogen). For hybridizations, we used human 4 × 44k whole-genome human arrays (Agilent). Isolation and identification of RNA-binding proteins were based on previously published protocols^{13,24}. HNRPA2B1 target transcripts were isolated based on the CLIP protocol¹⁸.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 August 2011; accepted 2 March 2012.

Published online 8 April 2012.

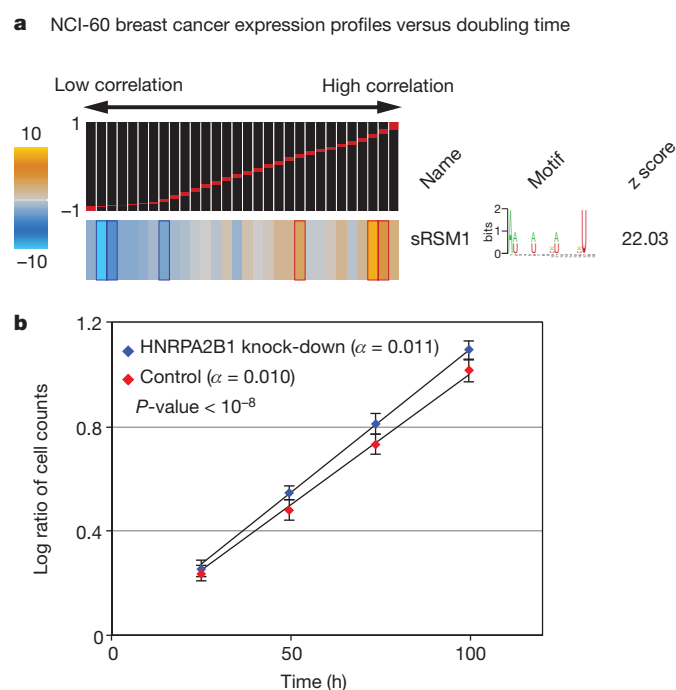


Figure 4 | HNRPA2B1 regulates growth rate. **a**, Whole genome expression levels across five breast cancer cell lines (MCF7, MDA-MB-231, HS578T, BT-549 and T47D) were correlated against their doubling times¹⁷. The resulting values, ranging from -1 to 1, were analysed by TEISER to probe the enrichment/depletion pattern of transcripts carrying sRSM1. **b**, The growth of HNRPA2B1 siRNA-transfected samples was compared to those of mock-transfected controls. For each time-point, the number of cells in four independent samples was counted in duplicates ($n = 8$), yielding an estimated growth-rate (α). Shown are the average log-ratios, their standard deviation at each time-point, and the statistical significance of the observed difference in growth-rate.

1. Dölken, L. *et al.* High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA* **14**, 1959–1972 (2008).
2. Elemento, O., Slonim, N. & Tavazoie, S. A universal framework for regulatory element discovery across all genomes and data types. *Mol. Cell* **28**, 337–350 (2007).
3. Rabani, M., Kertesz, M. & Segal, E. Computational prediction of RNA structural motifs involved in posttranscriptional regulatory processes. *Proc. Natl Acad. Sci. USA* **105**, 14885–14890 (2008).
4. Barash, Y. *et al.* Deciphering the splicing code. *Nature* **465**, 53–59 (2010).
5. Wan, Y., Kertesz, M., Spitale, R. C., Segal, E. & Chang, H. Y. Understanding the transcriptome through RNA structure. *Nature Rev. Genet.* **12**, 641–655 (2011).
6. Pavesi, G., Mauri, G., Stefani, M. & Pesole, G. RNAProfile: an algorithm for finding conserved secondary structure motifs in unaligned RNA sequences. *Nucleic Acids Res.* **32**, 3258–3269 (2004).
7. Searls, D. B. The language of genes. *Nature* **420**, 211–217 (2002).
8. Hofacker, I. L., Fekete, M. & Stadler, P. F. Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.* **319**, 1059–1066 (2002).
9. Kertesz, M. *et al.* Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467**, 103–107 (2010).
10. Goodarzi, H., Elemento, O. & Tavazoie, S. Revealing global regulatory perturbations across human cancers. *Mol. Cell* **36**, 900–911 (2009).
11. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
12. Cutroneo, K. R. & Ehrlich, H. Silencing or knocking out eukaryotic gene expression by oligodeoxynucleotide decoys. *Crit. Rev. Eukaryot. Gene Expr.* **16**, 23–30 (2006).

13. Windbichler, N. & Schroeder, R. Isolation of specific RNA-binding proteins using the streptomycin-binding RNA aptamer. *Nature Protocols* **1**, 637–640 (2006).
14. Biamonti, G., Ruggiu, M., Saccone, S., Della Valle, G. & Riva, S. Two homologous genes, originated by duplication, encode the human hnRNP proteins A2 and A1. *Nucleic Acids Res.* **22**, 1996–2002 (1994).
15. Wilusz, C. J., Wormington, M. & Peltz, S. W. The cap-to-tail guide to mRNA turnover. *Nature Rev. Mol. Cell Biol.* **2**, 237–246 (2001).
16. Michlewski, G. & Caceres, J. F. Antagonistic role of hnRNP A1 and KSRP in the regulation of *let-7a* biogenesis. *Nature Struct. Mol. Biol.* **17**, 1011–1018 (2010).
17. Ross, D. T. *et al.* Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genet.* **24**, 227–235 (2000).
18. Jensen, K. B. & Darnell, R. B. CLIP: crosslinking and immunoprecipitation of *in vivo* RNA targets of RNA-binding proteins. *Methods Mol. Biol.* **488**, 85–98 (2008).
19. Keene, J. D., Komisarow, J. M. & Friedersdorf, M. B. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nature Protocols* **1**, 302–307 (2006).
20. Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469 (2008).
21. Giannopoulou, E. G. & Elemento, O. An integrated ChIP-seq analysis platform with customizable workflows. *BMC Bioinformatics* **12**, 277–294 (2011).
22. Beer, M. A. & Tavazoie, S. Predicting gene expression from sequence. *Cell* **117**, 185–198 (2004).
23. Yang, Y. *et al.* RNA secondary structure in mutually exclusive splicing. *Nature Struct. Mol. Biol.* **18**, 159–168 (2011).
24. Greco, T. M., Yu, F., Guise, A. J. & Cristea, I. M. Nuclear import of histone deacetylase 5 by requisite nuclear localization signal phosphorylation. *Mol. Cell Proteomics* **10**, M110.004317 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the members of the Tavazoie laboratory for comments on the project and manuscript. We are also grateful to N. Pencheva, B. Tsui, S. Tavazoie and L. Dölken for their intellectual and technical contributions. L.F. was supported by a Ruth L. Kirschstein National Research Service Award (T32-GM066699). S.T. was supported by grants from NHGRI (2R01HG003219) and the NIH Director's Pioneer Award.

Author Contributions H.G., H.S.N. and S.T. conceived and designed the study. H.G. and H.S.N. developed TEISER. R.S. contributed to the execution of the study. H.G., H.S.N., T.M.G., P.O., I.M.C. and S.T. designed the experiments. H.G., P.O., L.F. and T.M.G. performed the experiments. H.G., H.S.N. and T.M.G. analysed the results. H.G., H.S.N. and S.T. wrote the paper.

Author Information The microarray and high-throughput sequencing data are deposited at GEO under the umbrella accession number GSE35800. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.T. (st2744@columbia.edu).

METHODS

TEISER: detailed description of the algorithm. Genome profile. A genome profile is defined across the genes in the genome, where each gene is associated with a unique measurement. Whole-genome measurements, discrete or continuous, can be obtained from a variety of experimental or computational sources (for example, Supplementary Fig. 1).

Structural motif definition. Each structural motif is defined as a series of context-free statements that define the structure and sequence of the motif (Supplementary Fig. 2). A context-free grammar is a set of production rules that describes how phrases are made from their building blocks. Considering a structured RNA molecule as a phrase, its potential building blocks are the different base pairs and bulges. Loops can be considered as bulges that happen at the beginning of phrases. Also, internal loops can be considered as combination of left and right bulges in the middle of phrases. The context-free grammar that we have used contains the following production rules: $S \rightarrow S[AUCGN]$, $S \rightarrow [AUCGN]S$, $S \rightarrow [AUCGN]S[AUCGN]$; wherein the first production rule depicts a right bulge, the second production rule results in a left bulge, and the third production rule creates a base-pairing. For example, consider the stem loop AAACGCUUU (the stem region is underlined). Let the symbol S be a non-terminal symbol that stands for this stem loop; the production rule $S \rightarrow SG$ adds a G to the 3' end of the molecule, creating a new S , AAACGCUUUG, which has an unpaired 3'-end G . Next, using the production rule $S \rightarrow GSC$, we can add a G to the 5' end and a C to the 3' end of the molecule and make them pair with each other, again creating a new S , GAAACGCUUUGC, which can be further expanded in this way. Note that the G that we added in the previous step has now become a right bulge.

Motif profile. For every given motif, we create a binary vector across all the genes, in which '1' denotes the presence and '0' denotes the absence of that motif. This vector is called a motif profile.

Creating seed CFGs. We used, as the seed motifs, an exhaustive set of context-free statements that represented all possible stem-loop structures that satisfied the following criteria: stem length of at least 4 bp and at most 7 bp; loop length of at least 4 nt and at most 9 nt; at least 4 and at most 6 production rules representing non-degenerate bases (that is, production rules that are not $S \rightarrow SN$, $S \rightarrow NS$, or $S \rightarrow NSN$); and information content of at least 14 bits and at most 20 bits. The information content of the motif M , which is represented by n production rules, was defined as $-\log_2(p_M)$, wherein p_M is the probability that a random sequence of length l matches the n production rules of motif M , with l being equal to $2 \times n_1 + n_2$ in which n_1 is the number of production rules that represent base pairings and n_2 is the number of production rules that represent bulges ($n_1 + n_2 = n$).

Quantizing continuous genome profiles. Mutual information is defined for both continuous and discrete random variables; however, in practice, continuous data are discretized before calculating the mutual information (MI) values. Our quantization procedure involves using equally populated 'bins'. Thus, the discretization step only requires a single parameter, that is, the number of genes in each bin. In TEISER, we have set the default number of bins to 30 ($N_e = 30$). It should be noted that the results are not sensitive to variations in the value of N_e as long as N_e is >10 and each bin has more than ~ 100 associated transcripts.

Removing recently duplicated genes. Recently duplicated members of gene families or transposons often share a significant amount of sequence identity in their UTRs. They also tend to cross-hybridize on the arrays and show a high artificial correlation. This would in turn bias our search towards conserved elements in the UTRs of these genes. In TEISER, similar to FIRE², we remove the duplicates that have similar values (for example, fall in the same bin after quantization of the input genome profile). A MegaBlast E -value cutoff of 1×10^{-15} was used to identify duplicates.

Calculating the mutual information values. We performed mutual information (MI) calculations between the genome profile and the motif profiles using algorithms introduced and described elsewhere^{2,10}. These algorithms take the necessary steps to ensure reliable MI calculations (for example, minimum sample sizes for reliable estimation of joint distributions).

Randomization-based statistical testing. To assess the statistical significance of the calculated MI values, TEISER uses a non-parametric randomization-based statistical test. In this test, the genome profile is shuffled 1,500,000 times and the corresponding MI values are calculated. A motif is deemed significant only if the real MI value is greater than all of the randomly generated ones. In TEISER, in order to minimize the required number of tests, structural motifs are first sorted based on the MI values (from high to low) and the statistical test is applied in order. When 20 contiguous motifs in the sorted list do not pass the test, the procedure is terminated.

Optimization of the identified seeds into more informative motifs. Our initial collection of structural motifs, despite being large, is a coarse-grained sampling of the entire space. Mainly, it provides us with a set of informative seeds that should be later optimized into closer representations of their actual form².

Accordingly, all the structural motifs that pass the previous stage are further optimized and elongated. The process involves: (1) optimization: randomly choose one of the context-free statements (production rules) from the motif and convert its sequence information to all possible combinations of nucleotides. Evaluate all the resulting structural motifs and accept the one that results in the highest MI value. (2) Elongation: production rules are added to the end of the context-free phrase that represents the motif, thus extending its effective length in the form of a base pair or a bulge. The increase in length is similarly accepted only if it results in a higher MI value.

Removing redundantly informative structural motifs. Motifs that redundantly represent the same potential *cis*-regulatory elements are identified and removed using the concept of conditional information as described before^{2,10}.

Finding robust motifs. TEISER also performs jack-knife resampling to find robust motifs that are not over-sensitive to the composition of the input data. For each predicted motif, we perform 10 jackknifing trials where, in each trial, one third of the genes are randomly removed and the mutual information value and its statistical significance is evaluated. The robustness score is then defined as the number of trials in which the motif remains significant (scores better in the original genome profile than in all the randomly shuffled genome profiles) after resampling, ranging from 0/10 to 10/10. By default, TEISER requires the motif to be significant in more than half of the trials (a robustness score equal to or greater than 6/10). While this parameter can be changed at the user's discretion, our experience with both TEISER and FIRE² suggests that this threshold results in very low false discovery rates across a variety of data sets (discrete and continuous). **Patterns of motif enrichment and depletion.** For a given motif, a high mutual information value results from the non-random distribution of its targets across the input range. This results in significant patterns of enrichment and depletions across the genome profile, which can be quantified by calculating enrichment/depletion scores. These scores result from the log transformation of P -values calculated based on the hypergeometric distribution, as described previously².

Final statistical tests. In case the genome profile is continuous, one can require TEISER to return motifs that are enriched at one end of the data range or the other (for example, structural motifs in Fig. 1). TEISER accomplishes this through calculating the Spearman correlation between the enrichment scores and the average data value across all the bins. For the structural motifs in Fig. 1, the P -value threshold for these Spearman correlations was set to 0.001 (for Supplementary Fig. 3, this value is 0.01 which puts the FDR at 10%). It should be noted, however, that other statistical tests could be used in this step at the discretion of the user. The goal, ultimately, is to identify the motifs that show significant enrichments at either end of the data range.

Inter-species conservation. For each motif, we also calculate a conservation score based on its network-level conservation with respect to a related genome². For this, orthologous transcripts in both genomes are scanned for the presence/absence of the motif. The overlap of positive sequences between the orthologous sequences is used to calculate a hypergeometric P -value². The conservation score is then defined as $1 - P$, which ranges between 0 and 1 (1 being highly conserved between the two genomes). In this study, we have used the human and mouse genomes to calculate the conservation scores associated with each structural motif.

Finding potentially active instances of each motif. As described previously², we defined the target genes of a predicted motif as all transcripts whose 3' or 5' UTRs contain the motif and are associated with a category/bin where the motif is enriched. In other words, these are the transcripts whose UTRs contain potentially 'active' motif occurrences. Upon identifying these likely targets for each structural motif, a weight-matrix can be generated from these potentially functional instances as a post-processing step (Supplementary Table 2).

False-discovery rate. In order to assess the false discovery rate, we ran 30 trials with shuffled 5' and 3' UTR sequences. In all the trials, not a single motif passed all the statistical tests. Thus, in case of the stability data set, the number of false positives in each trial, on average, is smaller than $1/30 \approx 0.34$, which corresponds to an FDR of <0.01 .

Predicting functional interactions. Given two motifs, structural or linear, one can assess their putative functional interaction through measuring how informative the presence of one would be about the presence or absence of the other. For revealing these interactions, we again use mutual information values calculated for pairwise motif profiles of structural and linear motifs. Randomization-based statistical tests are then used to find the significant interactions. For this, one of the motif profiles is shuffled 10,000 times and the interaction is deemed significant only if the real mutual information value is higher than all the 10,000 random ones. Predicting the target pathways. iPAGE¹⁰, with default settings, was used to identify the likely pathways that are regulated by the discovered structural and linear motifs.

Availability. TEISER is available online for download at <https://tavazoielab.c2b2.columbia.edu/TEISER>.

Measuring mRNA stability. RNA stability measurements were performed based on a previously published protocol¹. In short, MDA-MB-231 cells at 70% confluency were incubated in the presence of 25 μ M 4-thiouridine (Sigma) for 4 h. Then the cells were washed with fresh media (DMEM + 10% FBS) and incubated for 0, 1, 2 and 4 h. At each time point, cells were washed with cold PBS and RNA extraction was performed using a total RNA purification kit (Norgen Biotek). The 4-thiouridine thiol groups were then biotinylated using EZ-Link Biotin-HPDP (Pierce). We subsequently used μ Mac magnetic columns (Miltenyi Biotec) to capture the labelled RNAs. The resulting samples were then processed for one-colour hybridization using a one-colour low-input quick-amp labelling kit (Agilent) and hybridized according to the manufacture's instructions. A one-colour RNA spike-in kit (Agilent) was used as endogenous control to normalize values between arrays. For each transcript, the drop in signal as a function of time was used as a measure of mRNA stability (Supplementary Fig. 1): $r = -\ln\left(\frac{S_t}{S_0}\right) / t$, where S_t denotes signal at time t . Linear regression was used to calculate r for each transcript based on the hybridization signals from the four time points. It should be noted that TEISER is a non-parametric approach, thus it is the ranking rather than the actual stability values that underlies our motif discovery.

Transfection of decoy and scrambled oligonucleotides. We chose real instances of the sRSM1 structural motifs from NM_014363, which contains four instances of sRSM1, to create two decoy sets of sequences, each containing two of these instances (underlined) along with part of the real sequences as context. Set 1: AAAACTATTTTGAAGATGGTGGTGAGCTGCAAAATAGCTGGATGGATT TGAATGATTGGGATGATACATCATTGAACTGCACTTTATATAACCAAA GCTTAGCAGTTTGTAGATAAGAGTCTATGTATGTCCTCGTTAGGATG AAGTTAATTTTATGTTTTTAACATGGTATTTTGAAGGAGCTAATGAAA CACTGG. Set 2: ATTGTTTCTGAAACTGCTTGCCAAGACAACATTTATTA ACTGTTAGAACACCTTCTTTATGTTTGTGTACATATTTCCACAAAT GTTATAATTTATATAGTGTGGTTGAACAGGATGCAATCTTTTGTGTCT AAAGGTGCTGCAGTTAAAAAAAACCAACCTTTCTTTCAATATGGCAT GTAGTGGAGTTTTT. For the scrambled controls, we used the shuffled version of the putative binding sites (see Supplementary Fig. 5). These two decoy/scrambled sets were then chemically synthesized (IDT). An upstream T7 promoter was used to transcribe the constructs *in vitro* using Megascript T7 kit (Ambion). In order to reduce cytotoxicity, RNA molecules were capped and poly-A tailed using Cap Analogue (Ambion) and poly-A polymerase (NEB). MDA-MB-231 cells at 80% confluency were transfected with the resulting RNA oligos using Lipofectamin 2000 reagent (Invitrogen) according to manufacturer's recommendations. Experiments were performed in duplicates for each set. Forty-eight hours post-transfection, we extracted RNA and differentially labelled the samples with Cy3 or Cy5 dyes. The samples were then hybridized on Agilent human gene expression arrays (4 \times 44k). The Cy3/Cy5 ratios from the two biological replicates were then averaged into a single data set as log of ratios, which was then analysed by TEISER.

Reporter system for testing the functionality of sRSM1 instances. The plasmid pcDNA5/FRT/TOPO (Invitrogen) was used to clone a GFP-coding sequence along with a gateway cloning site downstream of GFP (in its 3' UTR). Decoy and scrambled sequences (Set 1 in the previous section) were subsequently cloned into the resulting construct using the gateway site. The resulting plasmids were transfected into the Flp-In 293 cell line (Invitrogen), and the cells were grown in Hygromycin for selecting stably transfected cells. The resulting cell lines, named Flp-In 293 GFP-Decoy and Flp-In 293 GFP-Shuffled, were subjected to FACS measurements to quantify GFP expression. For the decay rate measurements, cells were incubated in media with 5 μ g ml⁻¹ of α -amanitin (Sigma). Time points were taken at 0, 1.5, 3 and 6 h in duplicates for Flp-In 293 GFP-Decoy and Flp-In 293 GFP-Shuffled cells. Quantitative PCR (Fast SYBR Green Master Mix, Ambion) was then used to determine the relative quantity of GFP transcript in each cell line at different time-points using 18S rRNA as endogenous control.

Identifying binding candidates of sRSM1. We used a published protocol¹³ to isolate potential RNA-binding proteins that bind sRSM1. In short, the StreptoTag aptamer was added downstream of the Set 1 decoy and scrambled sequences. The resulting RNAs were then immobilized on a dihydrostreptomycin Sepharose column (GE Healthcare) and were used to immunoprecipitate potential partners. Total protein was extracted from MDA-MB-231 cells (Total Protein Extraction Kit, Millipore), 1,000 μ g of which was used as input to each column. Samples were then washed, eluted in 10 μ M streptomycin and subjected to in-solution digestion^{24,25}. Tryptic peptides were then analysed by nanoliquid chromatography-tandem mass spectrometry using an Ultimate 3000 nRSLC (Dionex) coupled online to an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific), as previously described²⁴.

HNRPA2B1 knock-down. The ON-Targetplus (Dharmacon) set of siRNAs for HNRPA2B1 (target sequences: GAGGAGGAUCUGAUGGAUA, GGAGAGUA GUUGAGCCAAA, and GCUGUUUGUUGGCGGAAUU) were used to transfect MDA-MB-231 cells (grown in D10F medium) using Lipofectamine 2000 (Invitrogen). Three of the four tested siRNAs resulted in a substantial knock-down in HNRPA2B1 (more than twofold reduction in expression, log ratio >0.4 and $P < 10^{-7}$) and their corresponding samples were used for hybridization. Forty-eight hours post-transfection, we extracted total RNA from each sample along with mock-transfected controls. We then differentially labelled the RNA samples with Cy3 and Cy5 dyes and hybridized them to Agilent human gene expression arrays (4 \times 44k). The log of signal ratios was used as a measure of differential expression between the samples and controls. These values were averaged across the three samples and were subsequently analysed by TEISER to assess the enrichment/depletion pattern of sRSM1 across the distribution.

For the decay rate measurements, forty-eight hours post-transfection, cells were incubated in media with 5 μ g ml⁻¹ of α -amanitin (Sigma). Time points were taken at 0, 1, 2 and 4 h in duplicates for the siRNA-transfected samples and mock-transfected controls. Each sample was then Cy3-labelled and hybridized to expression arrays (Agilent 4 \times 44k) in duplicates and the reported signals were used to calculate decay rates. Following this procedure, for each transcript, four decay rates (two biological replicates, each having two technical replicates) were calculated from the siRNA-transfected samples and four decay rates from the controls. For each transcript, we then calculated a value according to $s(1 - P)$, where P is the t -test P -value between the two sets and s denotes whether the decay rates are higher in the siRNA samples (+1) or the mock controls (-1). After this transformation, the data range is between -1 and 1 with the background genes (the transcripts that show little change between the two samples) around 0. TEISER was then used to visualize the enrichment pattern of sRSM1 across this data range.

Identifying transcripts that interact with HNRPA2B1 (RIP-chip). A myc-tagged ORF clone of HNRPA2B1 (variant A2, OriGene) was transfected into MDA-MB-231 cells (grown in D10F medium) using Lipofectamine LTX and Plus reagent (Invitrogen). Seventy-two hours post-transfection, the cells were washed with cold PBS and ultraviolet-irradiated at 4,000 mJ cm⁻². The cells were then collected and lysed with 1 ml M-PER Reagent (Pierce) and 10 μ l RNasin (NEB). The samples were subjected to DNase treatment (baseline ZERO DNase) for 15 min at 37 °C. Samples were then centrifuged at 16,000g at 4 °C for 20 min to pellet the cell debris. Immunoprecipitation of tagged HNRPA2B1 protein was performed using Mammalian c-Myc Tag IP/Co-IP Kit (Pierce) per manufacturer's instructions. Upon elution, samples were subjected to proteinase K digestion and polyadenylation. The RNA molecules in each sample were extracted using RNeasy MinElute Cleanup Kit (Qiagen) and Cy3-labelled using low-input quick-amp labelling kit (Agilent). As control, we used Cy5-labelled RNA samples extracted before HNRPA2B1 immunoprecipitation. The samples were hybridized to Agilent human gene expression arrays (4 \times 44k) and the log of signal ratios was used as a measure of transcript affinity to HNRPA2B1. For each transcript, affinity values were averaged across two biological replicates and TEISER was used to assess the enrichment/depletion pattern of sRSM1.

Identifying 3'UTR binding sites of HNRPA2B1 (HITS-CLIP). A strategy similar to that of target transcript identification was used to discover the HNRPA2B1 binding sites. Upon ultraviolet-irradiation of mycHNRPA2B1-transfected cells, the samples were subjected to the HITS-CLIP protocol previously described elsewhere²⁶. ChIPSeeq²¹, an integrated ChIP-seq analysis platform, was used to identify binding sites and extract real and random sequences (default parameters) for analysis with TEISER.

Measuring growth-rates in HNRPA2B1 knock-down cells. HNRPA2B1 siRNAs (Dharmacon) were used to knock-down the expression of this regulator. Seventy-two hours post-transfection, four independent samples were harvested and counted in duplicates as the baseline number of cells at time zero. Similarly, samples were counted at 25, 49.5, 73.5 and 99.5 h time-points. The same experiment was performed for mock-transfected cells. Using an exponential growth model, the log-ratio of the counted cells at each time-point was used to estimate a growth rate for siRNA-transfected and mock-transfected samples. ANCOVA was used to determine the P -value associated with the observed differences between the two growth rates.

25. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359–362 (2009).
26. Chi, S. W., Zang, J. B., Mele, A. & Darnell, R. B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).

Thermal and electrical conductivity of iron at Earth's core conditions

Monica Pozzo¹, Chris Davies², David Gubbins^{2,3} & Dario Alfè^{1,4}

The Earth acts as a gigantic heat engine driven by the decay of radiogenic isotopes and slow cooling, which gives rise to plate tectonics, volcanoes and mountain building. Another key product is the geomagnetic field, generated in the liquid iron core by a dynamo running on heat released by cooling and freezing (as the solid inner core grows), and on chemical convection (due to light elements expelled from the liquid on freezing). The power supplied to the geodynamo, measured by the heat flux across the core–mantle boundary (CMB), places constraints on Earth's evolution¹. Estimates of CMB heat flux^{2–5} depend on properties of iron mixtures under the extreme pressure and temperature conditions in the core, most critically on the thermal and electrical conductivities. These quantities remain poorly known because of inherent experimental and theoretical difficulties. Here we use density functional theory to compute these conductivities in liquid iron mixtures at core conditions from first principles—unlike previous estimates, which relied on extrapolations. The mixtures of iron, oxygen, sulphur and silicon are taken from earlier work⁶ and fit the seismologically determined core density and inner-core boundary density jump^{7,8}. We find both conductivities to be two to three times higher than estimates in current use. The changes are so large that core thermal histories and power requirements need to be reassessed. New estimates indicate that the adiabatic heat flux is 15 to 16 terawatts at the CMB, higher than present estimates of CMB heat flux based on mantle convection¹; the top of the core must be thermally stratified and any convection in the upper core must be driven by chemical convection against the adverse thermal buoyancy or lateral variations in CMB heat flow. Power for the geodynamo is greatly restricted, and future models of mantle evolution will need to incorporate a high CMB heat flux and explain the recent formation of the inner core.

First principles calculations of transport properties based on density functional theory (DFT) have been used in the past for a number of materials (see, for example, refs 9, 10). Recently, increased computer power has facilitated simulations of large systems, allowing the problem of the size of the simulation cell to be addressed: this can be a serious problem for the electrical conductivity, σ (ref. 11). Here we report a series of calculations of the electrical and thermal conductivity (k) of iron at Earth's core conditions, using DFT. We previously used these methods to compute an extensive number of thermodynamic properties of iron and iron alloys, including the whole melting curve of iron in the pressure range 50–400 GPa (refs 12, 13) and the chemical potentials of oxygen, sulphur and silicon in solid and liquid iron at inner core boundary (ICB) conditions, which we used to place constraints on core composition⁶. Recently, we computed the conductivity of iron at ambient conditions, and obtained values in very good agreement with experiments¹⁴.

We calculated three adiabatic temperature–pressure profiles (adiabats) for the core; to do this, we assumed three different possible temperatures at the ICB, and followed the line of constant entropy as the pressure was reduced to that of the CMB. The ICB temperatures

were: 6,350 K (the melting temperature of pure iron)¹³, 5,700 K (the melting temperature of a mixture of iron with 10% Si and 8% O, corresponding to an inner-core density jump $\Delta\rho = 0.6 \text{ g cm}^{-3}$)⁶ and 5,500 K (the melting temperature of a mixture of iron with 8% Si and 13% O, corresponding to $\Delta\rho = 0.8 \text{ g cm}^{-3}$)⁶. Then we calculated the electrical and thermal conductivity of iron at seven positions on these three adiabats. Our results are reported in Fig. 1, and show a smooth variation of these parameters in the core; σ only varies by $\sim 13\%$ between the ICB and the CMB, and it is almost the same for all adiabats. A recent shock wave experiment¹⁵ reported $\sigma = 0.765 \times 10^6 \Omega^{-1} \text{ m}^{-1}$ for pure iron at 208 GPa, and an older shock wave measurement¹⁶ reported $\sigma = 1.48 \times 10^6 \Omega^{-1} \text{ m}^{-1}$ at 140 GPa. Our values are closer to the latter. There is a larger variation in k , as implied by the Wiedemann–Franz law (which relates the thermal and electrical conductivity through $L = k/\sigma T$), which we found to be closely followed throughout the core with a Lorenz parameter $L = (2.48–2.5) \times 10^{-8} \text{ W } \Omega \text{ K}^{-2}$. The ionic contribution to k was calculated using the classical potential used as a reference system in ref. 12, which was shown to describe very accurately the energetics of the system and the structural and dynamical properties of liquid iron at Earth's core conditions. We found that the ionic contribution is only between 2.5 and $4 \text{ W m}^{-1} \text{ K}^{-1}$ on the adiabat, which is negligible compared to the electronic contribution, as expected.

The estimates of k (Fig. 1) are substantially larger than previously used in the geophysical literature, approximately doubling the heat conducted down the adiabatic gradient in the core and halving the power to drive a dynamo generating the same magnetic field. These considerations demand a revision of the power requirements for the geodynamo. The conductivities for liquid mixtures appropriate to the outer core are likely to be smaller than for pure iron, preliminary calculations suggesting about 30% lower, a smaller difference than that found in previous work¹⁷, but in close agreement with extrapolations obtained from recent diamond-anvil-cell experiments, which reported a value in the range $90–130 \text{ W m}^{-1} \text{ K}^{-1}$ at the top of the outer core¹⁸. Our values are also in broad agreement with recently reported DFT calculations¹⁹.

We focus on estimates for the two mixtures above, corresponding to ICB density jumps 0.6 g cm^{-3} (ref. 8) and 0.8 g cm^{-3} (ref. 7). There is relatively little effect on the conductivities in the two cases, because any additional O in the outer core must be balanced by less S or Si to maintain the mass of the whole core, which is well constrained. The larger density jump gives a higher O content, more gravitational energy, a lower ICB temperature and lower adiabatic gradient: it therefore favours compositional over thermal convection. The relevant values are given in Table 1.

We estimate power requirements for the dynamo using the model described in a previous study (ref. 5, and Methods). Neglecting small effects, the total CMB heat flux, Q_{CMB} , is the sum of terms proportional to either the CMB cooling rate, dT_0/dt , or the amount of radiogenic heating, h : $Q_{\text{CMB}} = Q_s + Q_L + Q_g + Q_r$, where the terms

¹Department of Earth Sciences, and Thomas Young Centre at UCL, UCL, Gower Street, London WC1E 6BT, UK. ²School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK. ³Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California at San Diego, 9500 Gilman Drive no. 0225, La Jolla, California 92093-0225, USA. ⁴Department of Physics and Astronomy, and London Centre for Nanotechnology, UCL, Gower Street, London WC1E 6BT, UK.

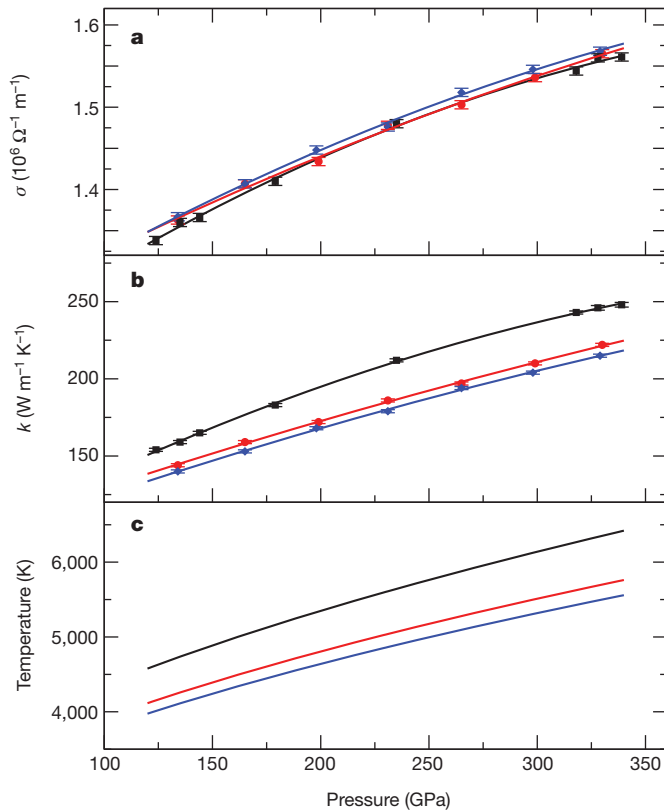


Figure 1 | Electrical and thermal conductivity of iron at Earth's outer core conditions. a–c, Electrical conductivity, σ (a), and electronic component of thermal conductivity, k (b), of pure iron corresponding to the three outer-core adiabatic profiles (adiabats) displayed in c. Black lines, adiabat corresponding to the melting temperature of pure iron at ICB pressure; red lines, that of the mixture containing 10% Si and 8% O; and blue lines, that of the mixture with 8% Si and 13% O. Lines are quadratic fits to the first principles raw data (symbols). Error bars (2 s.d.) are estimated from the scattering of the data obtained from 40 statistical independent configurations. Results are obtained with cells including 157 atoms and the single k -point (1/4,1/4,1/4), which are sufficient to obtain convergence within less than 1%.

on the right-hand side represent respectively the effects of secular cooling, latent heat, gravitational energy and radiogenic heating. The cooling rate, expressed in degrees per billion years, can be varied together with the radiogenic heating to produce some desired outcome: a fixed mantle heat flux, a marginal dynamo (no entropy left for ohmic dissipation, E_σ), or a primordial inner core (by decreasing the cooling rate and increasing the radiogenic heating). Results for a suite of 11 models are shown in Table 2.

Model 1 fails as a dynamo. There is an entropy deficit, meaning the assumption that the whole core can convect is incorrect—the temperature gradient must fall below the adiabat to balance the entropy equation. A dynamo might still be possible with a large part of the core completely stratified. Model 2 demonstrates the efficiency of compositional convection: the entropy is greatly increased compared to model 1 with no change in cooling rate and little increase in heat flux; the dynamo is now marginal. Model 3 has an increased cooling

Table 1 | Parameters used to estimate power requirements for the geodynamo

| $\Delta\rho$ | T_{ICB} | T_{CMB} | k_{ICB} | k_{CMB} | $\sigma_{\text{ICB}} (\times 10^6)$ | $\sigma_{\text{CMB}} (\times 10^6)$ | O | S/Si |
|--------------|------------------|------------------|------------------|------------------|-------------------------------------|-------------------------------------|----|------|
| 0.6 | 5,700 | 4,186 | 150 (223) | 100 (144) | 1.25 (1.56) | 1.11 (1.36) | 8 | 10 |
| 0.8 | 5,500 | 4,039 | 150 (215) | 100 (140) | 1.24 (1.57) | 1.11 (1.37) | 13 | 8 |

Values in parenthesis are for pure iron, other values are approximations for core mixtures. Units are g cm^{-3} for the ICB density jump, $\Delta\rho$; K for the temperatures, T ; $\text{W m}^{-1} \text{K}^{-1}$ for the thermal conductivity, k ; $\Omega^{-1} \text{m}^{-1}$ for the electrical conductivity, σ ; % for molar concentrations.

Table 2 | Heat flux and entropy for various models of cooling and radiogenic heating

| Model | $\Delta\rho$ | dT_{C}/dt | h | Q_{ad} | Q_{CMB} | IC age | E_σ | \mathcal{A} |
|-------|--------------|--------------------|-----|-----------------|------------------|--------|------------|---------------|
| 1 | 0.6 | 46 | 0 | 15.7 | 5.8 | 0.9 | -111 | 1,022 |
| 2 | 0.8 | 46 | 0 | 15.2 | 6.1 | 1.0 | 5 | 826 |
| 3 | 0.6 | 57 | 0 | 15.7 | 7.2 | 0.7 | -2 | 833 |
| 4 | 0.6 | 123 | 0 | 15.7 | 15.6 | 0.3 | 652 | 110 |
| 5 | 0.8 | 115 | 0 | 15.2 | 15.2 | 0.4 | 865 | 0 |
| 6 | 0.6 | 46 | 3.0 | 15.7 | 11.7 | 0.9 | 85 | 659 |
| 7 | 0.8 | 46 | 3.0 | 15.2 | 11.9 | 1.0 | 208 | 468 |
| 8 | 0.6 | 11.2 | 6.8 | 15.7 | 14.7 | 3.5 | -3 | 1,257 |
| 9 | 0.6 | 8.7 | 6.9 | 15.7 | 14.5 | 4.5 | -1 | 1,472 |
| 10 | 0.8 | 12.2 | 6.3 | 15.2 | 13.7 | 3.5 | 4 | 1,000 |
| 11 | 0.8 | 9.5 | 6.6 | 15.2 | 14.1 | 4.5 | 2 | 1,128 |

Here $\Delta\rho$ is the density jump at the ICB in g cm^{-3} ; dT_{C}/dt the cooling rate of the CMB in K Gyr^{-1} ; h the radiogenic heat source in pW kg^{-1} ; $Q_{\text{ad}} = -4\pi k(dT_{\text{ad}}/dr)$ is the heat conducted down the adiabat in TW where dT_{ad}/dr is the adiabatic gradient; Q_{CMB} is the heat flux across the CMB in TW; E_σ is the entropy available for the dynamo and other diffusive processes in MW K^{-1} . Inner core (IC) age is shown in Gyr; stable layer thicknesses, \mathcal{A} , are given in kilometres below the CMB.

rate and consequent younger inner core to demonstrate what is required for a marginal dynamo with $\Delta\rho = 0.6 \text{ g cm}^{-3}$. Models 4 and 5 have cooling rates that make the CMB thermally neutral; the CMB heat flux is equal to that conducted down the adiabat. Models 6 and 7 have some radiogenic heating and the original cooling rate and operate as dynamos, although they are still thermally stable at the top of the core. Models 8–11 have cooling rates that yield old inner-core ages, 3.5 and 4.5 Gyr, and the radiogenic heating has been adjusted to make a marginal dynamo. They are also thermally stable at the top of the core.

We estimate stable layer thicknesses by computing the radial variation of thermal and compositional gradients for each model using the equations of a previous study (ref. 20, Methods), which are derived from the equations of core energetics⁵. To compare thermal and chemical gradients, we multiply the latter by the ratio of compositional and thermal expansion coefficients α_c/α_T , thereby converting compositional effects into equivalent thermal effects. The base of the stable layer is defined as the point where the stabilizing adiabatic gradient, T'_a , crosses the combined destabilizing gradient, $T' = T'_L + T'_s + T'_c + T'_r$, where the terms represent respectively latent heat, secular cooling, compositional buoyancy and radiogenic heating.

Stable layer thicknesses are hundreds of kilometres in all models except those with cooling rates that are so rapid as to make the inner core too young; without compositional buoyancy the layers in all models except 4 and 5 span half the core (Table 2). Radiogenic heating thins the layers for the same cooling rate. Profiles of stabilizing and destabilizing gradients (Fig. 2) show that destabilizing gradients are greatest at depth, but much reduced compared to previous models²⁰ because they each depend on a factor $1/k$. The thermal conductivity increases by 50% across the core, increasing the heat conducted down the adiabat at depth and further reducing the power available to drive convection near the base of the outer core. Combined thermochemical profiles suggest that compositional buoyancy near the top of core is not strong enough to drive convection against the adverse temperature gradient.

Stable layers could be thinned or partially disturbed by convection, through penetration or instability, or some other effect not included in our simple model. A potentially more effective mechanism for inducing vertical mixing near the CMB is through lateral variations in CMB heat flux, which can drive motions without having to overcome the gravitational force. The presence of lateral variations makes the relevant heat flux for core mixing the maximum at the CMB²¹, which could be as much as 10 times the average²²; this does not influence dynamo entropy calculations but does allow magnetic flux to be carried to the surface in regions of cold mantle, as is observed²³.

As well as raising k , our calculations also raise σ to about twice the current estimate. Two important quantities depend on σ : the magnetic diffusion time (the time taken for the slowest decaying dipole mode to fall by a factor of e in the absence of a dynamo) and the magnetic Reynolds number R_m , which measures the rate of generation of

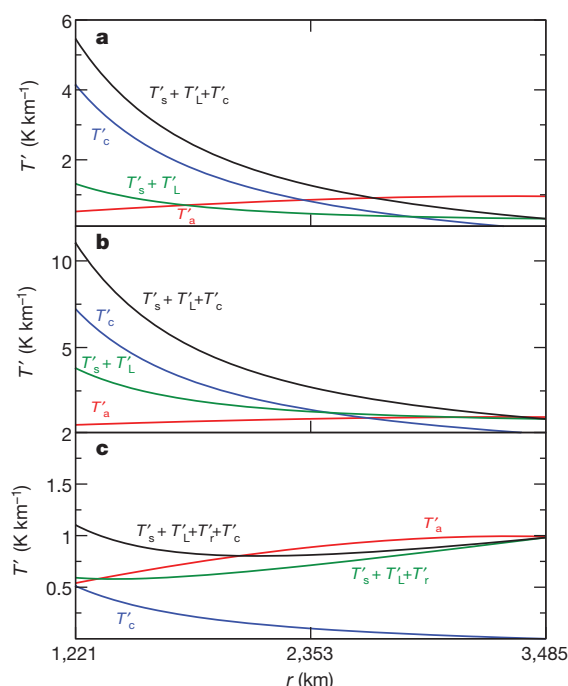


Figure 2 | Stabilizing and destabilizing gradients for three core energetics models. Equivalent temperature gradients, T' , plotted against radius for three core evolution models. The stabilizing gradient is due to conduction down the adiabat, T'_a (red lines). Compositional buoyancy is denoted by T'_c (blue lines), latent heat by T'_L , secular cooling by T'_s and radiogenic heating by T'_r . The total destabilizing thermal gradient is represented by the green lines; total destabilizing thermochemical gradients are represented by black lines. Three models from Table 2 are shown: **a**, model 2 ($\Delta\rho = 0.8 \text{ g cm}^{-3}$, $dT_o/dt = 46 \text{ K Gyr}^{-1}$ and $h = 0$); **b**, model 4 ($\Delta\rho = 0.6 \text{ g cm}^{-3}$, $dT_o/dt = 123 \text{ K Gyr}^{-1}$ and $h = 0$); **c**, model 9 ($\Delta\rho = 0.6 \text{ g cm}^{-3}$, $dT_o/dt = 8.7 \text{ K Gyr}^{-1}$ and $h = 6.9 \text{ pW kg}^{-1}$).

magnetic energy by a given flow. The magnetic diffusion time is increased to about 50 kyr. This may have significant implications for the theory of the secular variation: it makes the frozen flux approximation more accurate and lengthens the timescale of all diffusion-dominated processes, including polarity reversals. If current estimates of R_m are appropriate for the core²⁴, the increased conductivity implies that the geodynamo can operate on slower fluid flows and less input power from thermal and compositional convection.

Revised estimates of σ and k calculated directly at core conditions have fundamental consequences for the thermochemical evolution of the deep Earth. New estimates of the power requirements for the geodynamo suggest a CMB heat flux in the upper range of what is considered reasonable for mantle convection unless very marginal dynamo action can be sustained, while a primordial inner core is only possible with a significant concentration of radiogenic elements in the core. There are objections to a high CMB heat flux and also to radiogenic heating in the core^{25–27}, but one of the two seems inevitable if we are to have a dynamo. If the inner core is young, these high values of conductivity provide further problems with maintaining a purely thermally driven dynamo. A thermally stratified layer at the top of the core also appears inevitable. Viable thermal history models that produce thin stable layers and an inner core of age ~ 1 Gyr are likely to require a fairly rapid cooling rate and some radiogenic heating. The presence of a stable layer, and the effects associated with an increased electrical conductivity, have significant implications for our understanding of the geomagnetic secular variation.

METHODS SUMMARY

Calculations were performed using DFT with the same technical parameters used in refs 6, 12–14. We used the VASP code²⁸, PAW potentials^{29,30} with $4s^1 3d^7$ valence

configuration, the Perdew–Wang³¹ functional, a plane wave cut-off of 293 eV, and single particle orbitals were occupied according to Fermi–Dirac statistics. We tested the effect on the conductivity of the inclusion in valence of semi-core $3s$ and $3p$ states; we found that, as in the zero pressure case¹⁴, this effect is completely negligible.

The electrical conductivity and the electrical component of the thermal conductivity have been calculated using the Kubo–Greenwood formula and the Chester–Thellung–Kubo–Greenwood formula as implemented in VASP³². Because of the low mass of the electrons compared to the ions, the conductivities may be calculated by assuming frozen ionic configurations, and averaging over a sufficiently large set representing the typical distribution of the ions at the pressures and temperatures of interest.

Molecular dynamics simulations were performed in the canonical ensemble using cubic simulation cells with 157 atoms and the Γ point, a time step of 1 fs, and an efficient extrapolation of the charge density which speeds up the simulations by roughly a factor of two (ref. 33). Each state point was simulated for at least 6 ps, from which we discarded the first picosecond to allow for equilibration and used the last 5 ps to extract 40 configurations separated by 0.125 ps. This time interval is roughly two times longer than the correlation time, and therefore the configurations are statistically independent from each other. Because of the high temperatures involved, the conductivities converge quickly with respect to k -point sampling and size of the simulation cell¹⁴, and we found that with a 157-atom cells and the single k -point $(1/4, 1/4, 1/4)$ the results are converged to better than 1%.

The ionic component of the thermal conductivity was calculated using the Green–Kubo formula.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 November 2011; accepted 7 March 2012.

Published online 11 April 2012.

- Lay, T., Hernlund, J. & Buffett, B. Core-mantle boundary heat flow. *Nature Geosci.* **1**, 25–32 (2008).
- Labrosse, S., Poirier, J.-P. & Le Mouél, J.-L. On cooling of the Earth's core. *Phys. Earth Planet. Inter.* **99**, 1–17 (1997).
- Buffett, B., Garnero, E. & Jeanloz, R. Sediments at the top of Earth's core. *Science* **290**, 1338–1342 (2000).
- Lister, J. R. & Buffett, B. A. The strength and efficiency of thermal and compositional convection in the geodynamo. *Phys. Earth Planet. Inter.* **91**, 17–30 (1995).
- Gubbins, D., Alfè, D., Masters, T. G. & Price, D. Gross thermodynamics of 2-component core convection. *Geophys. J. Int.* **157**, 1407–1414 (2004).
- Alfè, D., Gillan, M. J. & Price, G. D. Temperature and composition of the Earth's core. *Contemp. Phys.* **48**, 63–80 (2007).
- Masters, T. G. & Gubbins, D. On the resolution of density within the Earth. *Phys. Earth Planet. Inter.* **140**, 159–167 (2003).
- Dziewonski, A. M. & Anderson, D. L. Preliminary Reference Earth Model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
- Silvestrelli, P. L., Alavi, A. & Parrinello, M. Electrical conductivity calculation in ab initio simulations of metals: application to liquid sodium. *Phys. Rev. B* **55**, 15515–15522 (1997).
- Mattsson, T. R. & Desjarlais, M. P. Phase diagram and electrical conductivity of high energy density water from density functional theory. *Phys. Rev. Lett.* **97**, 017801 (2007).
- Pozzo, M., Desjarlais, M. P. & Alfè, D. Electrical and thermal conductivity of liquid sodium from first principles calculations. *Phys. Rev. B* **84**, 054203 (2011).
- Alfè, D., Gillan, M. J. & Price, G. D. The melting curve of iron at the pressures of the Earth's core conditions. *Nature* **401**, 462–464 (1999).
- Alfè, D. Temperature of the inner-core boundary of the Earth: melting of iron at high pressure from first-principles coexistence simulations. *Phys. Rev. B* **79**, 060101(R) (2009).
- Alfè, D., Pozzo, M. & Desjarlais, M. P. Lattice electrical resistivity of magnetic body-centred cubic iron from first principles calculations. *Phys. Rev. B* **85**, 024102 (2012).
- Bi, Y., Tan, H. & Jing, F. Electrical conductivity of iron under shock compression up to 200 GPa. *J. Phys. Condens. Matter* **14**, 10849–10854 (2002).
- Keeler, R. N. & Royce, E. B. in *Physics of High Energy Density* (eds Caldirola, P. & Knoepfel, H.) 106–125 (Proc. Int. Sch. Phys. Enrico Fermi Vol. 48, 1971).
- Stacey, F. D. & Anderson, O. L. Electrical and thermal conductivities of Fe–Ni–Si alloy under core conditions. *Phys. Earth Planet. Inter.* **124**, 153–162 (2001).
- Hirose, K., Gomi, H., Ohta, K., Labrosse, S. & Hernlund, J. The high conductivity of iron and thermal evolution of the Earth's core. *Mineral. Mag.* **75**, 1027 (2011).
- de Koker, N., Steinle-Neumann, G. & Vlcek, V. Electrical resistivity and thermal conductivity of liquid Fe alloys at high P and T, and heat flux in Earth's core. *Proc. Natl Acad. Sci.* **109**, 4070–4073 (2012).
- Davies, C. J. & Gubbins, D. A buoyancy profile for the Earth's core. *Geophys. J. Int.* **187**, 549–563 (2011).
- Olson, P. in *Earth's Core and Lower Mantle* (eds Jones, C., Soward, A. & Zhang, K.) 1–49 (Taylor and Francis, London, 2000).

22. Nakagawa, T., & Tackley, P. J. Lateral variations in CMB heat flux and deep mantle seismic velocity caused by a thermal-chemical-phase boundary layer in 3D spherical convection. *Earth Planet. Sci. Lett.* **271**, 348–358 (2008).
23. Jackson, A., Jonkers, A. R. T. & Walker, M. R. Four centuries of geomagnetic secular variation from historical records. *Phil. Trans. R. Soc. Lond. B* **358**, 957–990 (2000).
24. Gubbins, D. in *Encyclopedia of Geomagnetism and Paleomagnetism* (eds Gubbins, D. & Herrero-Bervera, E.) 287–300 (Springer, 2007).
25. Davies, G. Topography: a robust constraint on mantle fluxes. *Chem. Geol.* **145**, 479–489 (1998).
26. Davies, G. Mantle regulation of core cooling: a geodynamo without core radioactivity? *Phys. Earth Planet. Inter.* **160**, 215–229 (2007).
27. McDonough, W. in *Treatise on Geochemistry* Vol. 2 (ed. Carlson, R. W.) 547–568 (Elsevier, 2003).
28. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
29. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
30. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
31. Wang, Y. & Perdew, J. P. Correlation hole of the spin-polarized electron gas, with exact small-wave-vector and high-density scaling. *Phys. Rev. B* **44**, 13298–13307 (1991).
32. Desjarlais, M. P., Kress, J. D. & Collins, L. A. Electrical conductivity for warm, dense aluminum plasmas and liquids. *Phys. Rev. E* **66**, 025401(R) (2002).
33. Alfè, D. Ab initio molecular dynamics, a simple algorithm for charge extrapolation. *Comput. Phys. Commun.* **118**, 31–33 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements D.G. is supported by CSEDI grant EAR1065597 from the National Science Foundation. C.D. is supported by a Natural Environment Research Council personal fellowship, NE/H01571X/1. M.P. is supported by NERC grant NE/H02462X/1 to D.A. Calculations were performed on the UK national facility HECToR.

Author Contributions D.A. and D.G. designed the project. M.P. and D.A. performed the first principles calculations. C.D. and D.G. performed the thermal history and core stratification calculations. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.A. (d.alf@ucl.ac.uk).

METHODS

First principles calculations. Calculations were performed using DFT with the same technical parameters used in refs 6, 12–14. We used the VASP code²⁸, PAW potentials^{29,30} with $4s^1 3d^7$ valence configuration, the Perdew–Wang³¹ functional, a plane wave cut-off of 293 eV, and single particle orbitals were occupied according to Fermi–Dirac statistics. We tested the effect on the conductivity of the inclusion in valence of semi-core $3s$ and $3p$ states; we found that, as in the zero pressure case¹⁴, this effect is completely negligible.

The electrical conductivity and the electrical component of the thermal conductivity have been calculated using the Kubo–Greenwood formula and the Chester–Thellung–Kubo–Greenwood formula as implemented in VASP³². Because of the low mass of the electrons compared to the ions, the conductivities may be calculated by assuming frozen ionic configurations, and averaging over a sufficiently large set representing the typical distribution of the ions at the pressures and temperatures of interest.

Molecular dynamics simulations were performed in the canonical ensemble using cubic simulation cells with 157 atoms and the Γ point, a time step of 1 fs, and an efficient extrapolation of the charge density which speeds up the simulations by roughly a factor of two (ref. 33). Each state point was simulated for at least 6 ps, from which we discarded the first picosecond to allow for equilibration and used the last 5 ps to extract 40 configurations separated by 0.125 ps. This time interval is roughly two times longer than the correlation time, and therefore the configurations are statistically independent from each other. Because of the high temperatures involved, the conductivities converge quickly with respect to k -point sampling and size of the simulation cell¹⁴, and we found that with a 157-atom cells and the single k -point $(1/4, 1/4, 1/4)$ the results are converged to better than 1%.

The ionic component of the thermal conductivity was calculated using the Green–Kubo formula.

Power estimates for the geodynamo. Estimates of the power required to drive the geodynamo are obtained by considering the slow evolution of the Earth using equations describing the balances of energy and entropy in the core. A detailed derivation of these equations can be found in a previous study⁵. Conservation of energy simply equates the heat crossing the CMB to the sources within: specific heat of cooling Q_s , latent heat of freezing Q_L , radiogenic heating Q_r , gravitational energy loss Q_g that is converted into heat by the frictional processes associated with the convection (almost entirely magnetic), and smaller terms⁵ involving pressure changes and chemistry that we shall ignore:

$$Q_{\text{CMB}} = Q_s + Q_L + Q_g + Q_r \quad (1)$$

All terms on the right-hand side of equation (1) can be written in terms of either the cooling rate at the CMB, dT_0/dt , or the amount of radiogenic heating, h . There is no dependence on the conductivities or the magnetic field, which are merely agents by which energy is converted to heat within the core.

These quantities do enter the entropy balance, however. This equation has dissipation terms from thermal and electrical conduction, plus viscosity and molecular diffusion. They are all positive because of the second law of thermodynamics. They are balanced by entropies associated with the power driving the convection: heat pumped in at a higher temperature and removed at a lower temperature (T_{CMB}) and gravitational energy that directly stirs the core and is converted to heat by frictional processes, the heat then being convected and conducted away. Note that entropy from heat is multiplied by a Carnot-like ‘efficiency factor’, $1/T_{\text{out}} - 1/T_{\text{in}}$ (latent heat is the most efficient because it is released at the highest temperature and removed at the lowest), while the gravitational energy is not, $E_g = Q_g/T_{\text{ICB}}$. Gravitational energy is more efficient at removing entropy and therefore more efficient than heat at generating magnetic field.

$$E = E_s + E_L + E_r + E_g = E_k + E_\sigma + E_\alpha \quad (2)$$

where the four terms on the left-hand side of the second equality represent secular cooling, latent heat release, radiogenic heating and gravitational energy loss. Adiabatic conduction entropy, E_k , is easily estimated from the thermal conductivity and adiabatic gradient and is large, of order 10^8 W K^{-1} . The new estimate of conductivity doubles older ones and the higher ICB temperatures increase it still further. Barodiffusion, E_α , is the tendency for light elements to migrate down a pressure gradient and its associated entropy is significant but small, not exceeding 2.5 MW K^{-1} in any of our estimates. Diffusional processes associated with convection and the geodynamo also produce entropy, denoted E_σ , mainly in the small scales. This presents a problem in estimation because the dominant contribution comes from magnetic fields, fluid flows, temperature and compositional fluctuations that cannot be observed and, in many cases, cannot even be simulated numerically. A low value of the power required to drive the dynamo, 0.5 TW

(ref. 34), was obtained from a numerical dynamo simulation³⁵, which at an average temperature of 5,000 K translates into $E_\sigma = 10^7 \text{ W K}^{-1}$, an order of magnitude lower than E_k , but the numerical simulation necessarily reduces small scale magnetic fields and the value for the Earth could be much larger. It may well be that future numerical simulations with higher resolution will have higher ohmic dissipation approaching E_k . Magnetic diffusivity is much larger than any other diffusivity in the core, by many orders of magnitude, and in numerical simulations the viscosity, thermal, and molecular diffusivities are replaced with turbulent values to account for unresolved, turbulent, small scale fields. Even so, the associated entropies remain much smaller than those associated with magnetic fields: they are generally ignored, although we should bear in mind that they are all positive and could make a contribution.

Parameter values used to calculate thermal contributions to the energy and entropy balances equations (1) and (2) are taken from Table 1 of a previous study³⁶, except for the thermal conductivity and the temperatures of the CMB and ICB, which are taken from the present study. Latent heat, Q_L , depends on τ , the difference between the melting and adiabatic gradients at the ICB; the value for the former is taken to be 9 K GPa^{-1} (ref. 36), while the value of the latter is calculated from Fig. 1 of this study. Parameter values used to calculate compositional terms differ slightly from previous work⁵, owing to their use of different concentrations for the light elements O, Si and S in the outer core. Concentration enters the calculation of gravitational energy through equation (9) of ref. 5, which, along with equation (8) of ref. 5, is used to define Q_g in equation (18) of ref. 5. Note also Q_g depends on τ . The remaining changes affect the barodiffusion, E_α , which makes a small contribution to the entropy budget (2); for completeness we list the new parameter values required to determine E_α in Supplementary Tables 1 and 2.

Estimating stable layer thicknesses. Radial profiles of the thermal and compositional energy sources that power the dynamo are determined using the equations of a previous study²⁰, which are derived from the energy balance appropriate for the outer core⁵. The radial profiles represent conductive solutions that satisfy the total CMB heat-flux boundary condition for the temperature, zero CMB mass flux of light elements, and fixed temperature and light element concentration at the ICB²⁰. Superimposed on this basic state are the small fluctuations associated with core convection and the dynamo process.

These radial profiles apply to a Boussinesq fluid and hence neglect compressibility effects other than when they act to modify gravity. This necessitates the use of an approximate form for the adiabatic temperature, a simple choice being a quadratic equation expressed in terms of the ICB and CMB temperatures¹⁹. Despite these simplifications, the CMB heat fluxes computed from equations (23)–(27) of the incompressible model²⁰ are in good agreement with those obtained from the original equations⁵ (see Supplementary Table 3), while the quadratic approximation for the adiabat differs by at most 10 K from the full calculation shown in Fig. 1.

Compositional buoyancy is at least as important for driving the geodynamo as thermal buoyancy (see, for example, ref. 5) and so we require a means of comparing the two in radial profiles, which is readily achieved by multiplying the former by the ratio of compositional and thermal expansion coefficients, α_c/α_T . This simple device converts compositional effects into equivalent thermal effects, thereby allowing all sources of buoyancy to be combined; it is also related to the condition of neutral stability discussed below. (However, it must be understood that the compositional term resulting from this transformation has nothing to do with the gravitational energy, Q_g , which is neglected in the Boussinesq equations³⁷.) We use the common approach (see, for example, ref. 37) of defining all fluxes that represent sources of buoyancy associated with the convection in terms of a turbulent diffusivity, which is assumed constant. By contrast, the heat flux due to conduction down the adiabatic gradient and the equivalent thermal flux due to barodiffusion must be defined in terms of molecular quantities.

The depth variation of the molecular thermal conductivity obtained from the DFT results is readily incorporated into the formulation of previous work²⁰. We write $k = k(r)$ to express the radial variation of the molecular thermal conductivity; equation (8) from ref. 20 must then be replaced by $q_a = \nabla \cdot (k(r) \nabla T_a)$, where ∇T_a is calculated from equation (12) in ref. 20. $k(r)$ is well-approximated by a parabolic conductivity variation, $k(r) = ar^2 + br + c$, which we use to calculate the heat flux down the adiabatic gradient.

To investigate the presence of a stable layer, we use temperature gradients instead of heat fluxes, which are calculated using equations (30)–(34) of a previous study²⁰ with $k(r)$ replacing k in the numerator of equation (30) of ref. 20. The parameter values are the same as those used to estimate power requirements above. We define the base of the stable layer to be the point of neutral stability as given by Schwarzschild’s criterion³⁸.

$$\left(\frac{dT}{dr} - \frac{dT_a}{dr}\right) + \frac{\alpha_c}{\alpha_T} \left(\frac{dc}{dr}\right) = 0$$

where dT/dr is the total temperature gradient, dT_a/dr is the adiabatic temperature gradient and dc/dr is the total compositional gradient. We write this condition as $T' = T'_L + T'_s + T'_c + T'_r - T'_a = 0$, where the terms represent respectively latent heat, secular cooling, compositional buoyancy, radiogenic heating and the adiabat, and prime indicates differentiation with respect to r (the barodiffusive contribution to dc/dr is very small and has been omitted). Possible deviations from the layer thicknesses we obtain using this definition can only be obtained by solving the complete dynamo equations with correct parameters for the Earth, which is

impossible at present. We believe this to be the best definition of the base of the layer given the nature of our thermodynamic model.

34. Buffett, B. A. Estimates of heat flow in the deep mantle based on the power requirements for the geodynamo. *Geophys. Res. Lett.* **29**, 1566–1569 (2002).
35. Kuang, W. & Bloxham, J. An Earth-like numerical dynamo model. *Nature* **389**, 371–374 (1997).
36. Gubbins, D., Alfè, D., Masters, T. G., Price, D. & Gillan, M. J. Can the Earth's dynamo run on heat alone? *Geophys. J. Int.* **155**, 609–622 (2003).
37. Anufriev, A. P., Jones, C. A. & Soward, A. M. The Boussinesq and anelastic liquid approximations for convection in the Earth's core. *Phys. Earth Planet. Inter.* **152**, 163–190 (2005).
38. Gubbins, D. & Roberts, P. H. in *Geomagnetism* (ed. Jacobs, J. A.) 30–32 (Academic, 1987).

Clusters of iron-rich cells in the upper beak of pigeons are macrophages not magnetosensitive neurons

Christoph Daniel Treiber¹, Marion Claudia Salzer¹, Johannes Riegler², Nathaniel Edelman¹, Cristina Sugar¹, Martin Breuss¹, Paul Pichler¹, Herve Cadiou³, Martin Saunders⁴, Mark Lythgoe², Jeremy Shaw⁴ & David Anthony Keays¹

Understanding the molecular and cellular mechanisms that mediate magnetosensation in vertebrates is a formidable scientific problem^{1,2}. One hypothesis is that magnetic information is transduced into neuronal impulses by using a magnetite-based magnetoreceptor^{3,4}. Previous studies claim to have identified a magnetic sense system in the pigeon, common to avian species, which consists of magnetite-containing trigeminal afferents located at six specific loci in the rostral subepidermis of the beak^{5–8}. These studies have been widely accepted in the field and heavily relied upon by both behavioural biologists and physicists^{9–11}. Here we show that clusters of iron-rich cells in the rostro-medial upper beak of the pigeon *Columbia livia* are macrophages, not magnetosensitive neurons. Our systematic characterization of the pigeon upper beak identified iron-rich cells in the stratum laxum of the subepidermis, the basal region of the respiratory epithelium and the apex of feather follicles. Using a three-dimensional blueprint of the pigeon beak created by magnetic resonance imaging and computed tomography, we mapped the location of iron-rich cells, revealing unexpected variation in their distribution and number—an observation that is inconsistent with a role in magnetic sensation. Ultrastructure analysis of these cells, which are not unique to the beak, showed that their subcellular architecture includes ferritin-like granules, siderosomes, haemosiderin and filopodia, characteristics of iron-rich macrophages. Our conclusion that these cells are macrophages and not magnetosensitive neurons is supported by immunohistological studies showing co-localization with the antigen-presenting molecule major histocompatibility complex class II. Our work necessitates a renewed search for the true magnetite-dependent magnetoreceptor in birds.

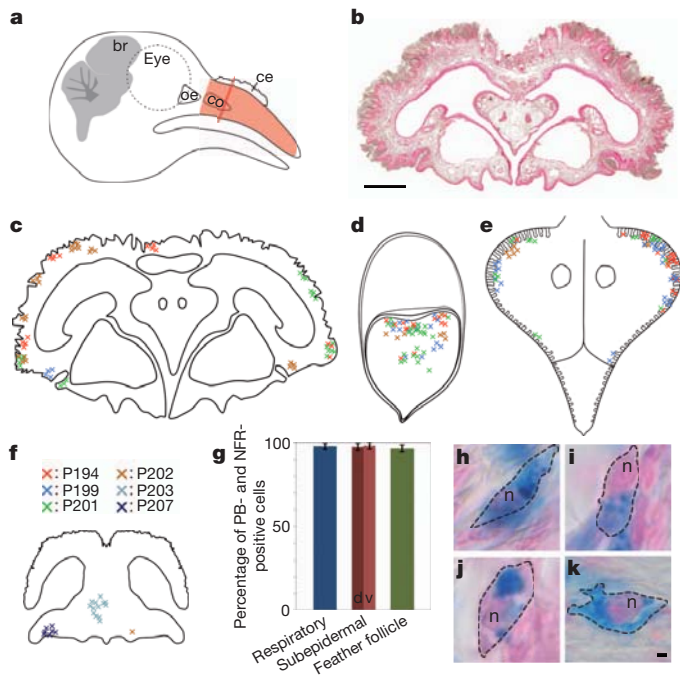
Each year millions of birds complete lengthy journeys guided by the Earth's magnetic field. Current evidence indicates that the detection of magnetic fields is mediated by an inclination-sensitive light-dependent compass that resides in the retina^{12,13}, and an intensity-sensitive apparatus that is believed to provide information about the magnetic map, is associated with the trigeminal nerve, and is thought to rely on biogenic magnetite (Fe₃O₄)¹⁴. The trigeminal nerve was first implicated in magnetoreception in the bobolink *Dolichonyx oryzivorus*, and was suggested to be sensitive to small alterations in magnetic stimuli¹⁵. Subsequent studies revealed that the ophthalmic branch is required for pigeons to perform an intensity-based conditioning task³, and that neurons in the trigeminal brainstem complex of European robins are activated when the birds are subjected to non-uniform magnetic fields¹⁶. These data have led to the proposition that the sensory cells responsible for magnetite-based magnetoreception lie in the upper beak of birds^{3,17}. Previous studies^{3,7} have claimed that clusters of iron-containing neurons in six specific bilateral locations in the rostral dermis of the upper beak of pigeons constitute a magnetic sense system⁷. It has been contended that this system consists of unmyelinated dendrites that contain superparamagnetic spherules surrounded by iron platelets that are composed of magnetite and

maghemite, and that the system is a common sensory apparatus in birds^{5–7}. These assertions have formed the basis for a host of behavioural studies and theoretical calculations that aim to advance the magnetite theory of magnetoreception^{9,10,18–20}.

To investigate this putative magnetic sense system, we undertook a systematic analysis of the prevalence and distribution of all iron-rich cells in the upper beak of the pigeon. We perfused adult pigeons (Nuremberg cohort, $n = 12$), and sectioned the upper beak from the caudal respiratory concha to the tip of the beak in the coronal plane (Fig. 1a, b). We stained serial sections (10 μ m) with Prussian blue (PB) to label ferric iron, and nuclear fast red (NFR) to identify nuclei, followed by counting of all PB-positive cells. We consistently observed PB-positive cells in three specific regions: (1) in the stratum laxum of the dorsal and/or ventral subepidermis; (2) in the buds of feather follicles; and (3) in the basal region of the respiratory epithelium (Fig. 1c–e). We confirmed this pattern of staining in a larger collection of pigeons originating from seven different lofts ($n = 172$). PB-positive cells in all three regions were characterized by the presence of multiple dark blue spherules (0.25–5.0 μ m in size) and/or by light blue cytoplasmic staining with a notable nucleus (Fig. 1g–k and Supplementary Figs 1–3). Subepidermal PB-positive cells in caudal and medial regions were predominantly found in the dorsal subepidermis (Fig. 1c), whereas those PB-positive cells located rostrally were found in the ventral subepidermis lining the inner roof of the beak (Fig. 1f and Supplementary Fig. 1). PB-positive cells in the feather follicle clustered in the apical region of the bud (Fig. 1d and Supplementary Fig. 2), and those in the respiratory epithelium were predominantly found within the lateral edges of the concha (Fig. 1e and Supplementary Fig. 3).

As it is believed that iron-rich cells in the upper pigeon beak are limited to six discrete bilateral anatomical loci⁵, we mapped the distribution of PB-positive cells along the rostro-caudal axis of the beak. To do this accurately we first created a three-dimensional topographic map of the pigeon beak by undertaking high-resolution magnetic resonance imaging (MRI) coupled with micro-computed tomography (micro-CT) scanning, identifying four specific anatomical landmarks (Supplementary Movies 1, 2 and Supplementary Figs 4, 5). After staining serial sections, we counted PB-positive cells and used our landmarks to map the distribution of cells along the rostro-caudal axis ($n = 12$). We found that PB-positive cells in the respiratory epithelium and feather follicles were restricted to caudal regions, whereas those in the subepidermal region were found in clusters along the length of the beak with no apparent bilateralization (Fig. 2a–c and Supplementary Figs 6, 7 and Supplementary Table 1). We found no significant differences in the total number of PB-positive cells between sexes (respiratory epithelium ($P > 0.5$), subepidermis ($P > 0.1$), feather follicle ($P > 0.1$)) (Supplementary Fig. 8), but observed an extremely large variation in the number and distribution of PB-positive cells when comparing birds of the same age and sex. For instance, pigeon 200 had ~200 PB-positive cells in the subepidermis, whereas pigeon

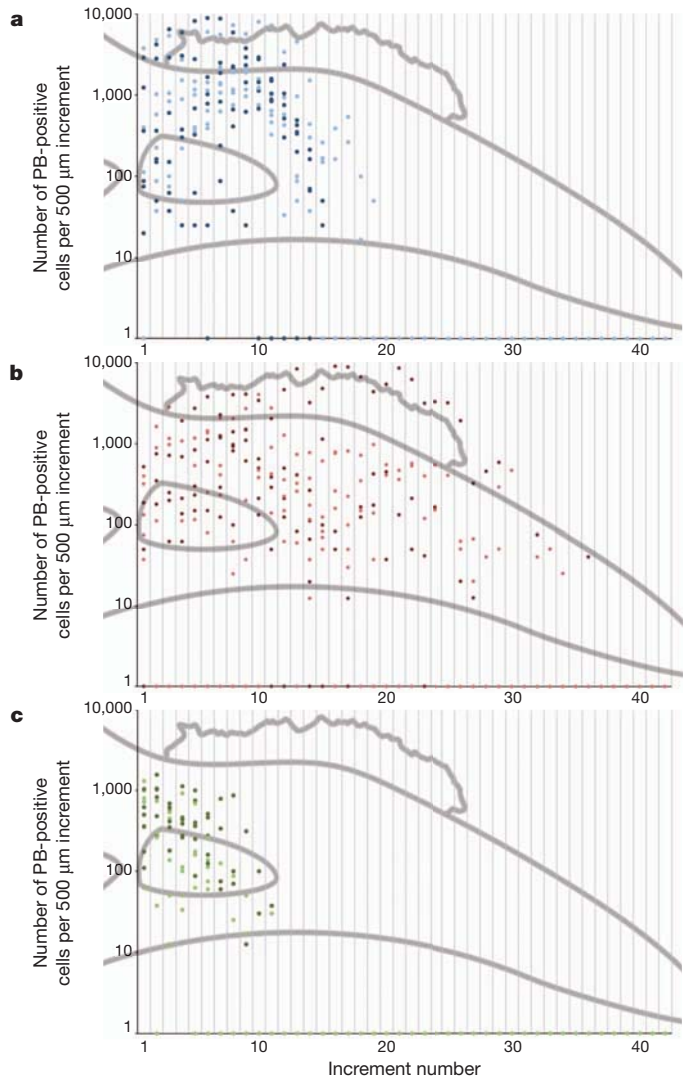
¹Institute of Molecular Pathology, Dr Bohr-Gasse, 1030 Vienna, Austria. ²Centre for Advanced Biomedical Imaging (CABI), Department of Medicine and Institute of Child Health, University College London (UCL), London WC1E 6DD, UK. ³Université de Strasbourg, Institut des Neurosciences Cellulaires et Intégratives (INCI), CNRS UPR 3212, F-67084 Strasbourg, France. ⁴Centre for Microscopy, Characterisation and Analysis, The University of Western Australia, Crawley 6009, Australia.



203 had ~108,800 PB-positive cells located in numerous clusters along the length of the beak. Although our serial quantification samples a section every 120 μ m we did not find the six 350- μ m-long bilateral clusters that are claimed to constitute a magnetic sense system⁷. We speculated that our pigeon strain might harbour a large genomic deletion that accounts for the absence of this putative magnetosensitive system. To investigate this we quantified PB-positive cells in pigeons from another loft (Vienna cohort, $n = 6$). Similar to our Nuremberg cohort we did not find six bilateral clusters, and once again observed a large variation in the distribution and number of PB-positive cells (Supplementary Fig. 9 and Supplementary Table 2). This variation is not consistent with a genetically encoded sensory apparatus responsible for magnetosensation.

Next we asked whether PB-positive cells are neurons by triple staining sections with PB, NFR, and one of three different antibodies that label neuronal structures: neurofilament (NF), TUBB3 and MAP1B ($n \geq 5$ birds). In the respiratory epithelium we observed 0.04% co-localization with NF ($n = 1,208$ cells), 0.6% co-localization with TUBB3 ($n = 2,818$ cells), and 0.01% co-localization with MAP1B ($n = 2,213$ cells). In the subepidermis we found no co-localization with NF ($n = 471$ cells) or MAP1B ($n = 803$ cells), and only 0.06% co-localization with TUBB3 ($n = 1,309$ cells). Finally, in the feather follicle we found no co-localization with NF ($n = 286$ cells) or MAP1B ($n = 295$ cells), and only 0.24% co-localization with TUBB3 ($n = 407$ cells) (Supplementary Figs 10, 11). The simplest explanation for the

very small amount of apparent co-localization we observed is that two cells, one PB positive and the other positive for a neuronal marker, lie in the same vertical plane, and because of the nature of the chemical stain used cannot be distinguished from one another. Taken together, our results strongly suggest that the clusters of PB-positive cells in the beak of the pigeon are not neurons.



To ascertain the true identity of the PB-positive cells, we undertook an analysis of their ultrastructure using transmission electron microscopy (TEM) ($n = 3$ birds) (Supplementary Fig. 12). We observed ferritin-like granules (6–9 nm) throughout the cytoplasm of PB-positive cells from all regions and in some instances haemosiderin masses and/or membrane-bound electron-dense organelles known as siderosomes^{21,22} (~300 nm) (Fig. 3a–f and Supplementary Fig. 13). Energy-filtered transmission electron microscopy (EFTEM) confirmed that each electron-dense granule was composed of iron (Supplementary

very small amount of apparent co-localization we observed is that two cells, one PB positive and the other positive for a neuronal marker, lie in the same vertical plane, and because of the nature of the chemical stain used cannot be distinguished from one another. Taken together, our results strongly suggest that the clusters of PB-positive cells in the beak of the pigeon are not neurons.

To ascertain the true identity of the PB-positive cells, we undertook an analysis of their ultrastructure using transmission electron microscopy (TEM) ($n = 3$ birds) (Supplementary Fig. 12). We observed ferritin-like granules (6–9 nm) throughout the cytoplasm of PB-positive cells from all regions and in some instances haemosiderin masses and/or membrane-bound electron-dense organelles known as siderosomes^{21,22} (~300 nm) (Fig. 3a–f and Supplementary Fig. 13). Energy-filtered transmission electron microscopy (EFTEM) confirmed that each electron-dense granule was composed of iron (Supplementary

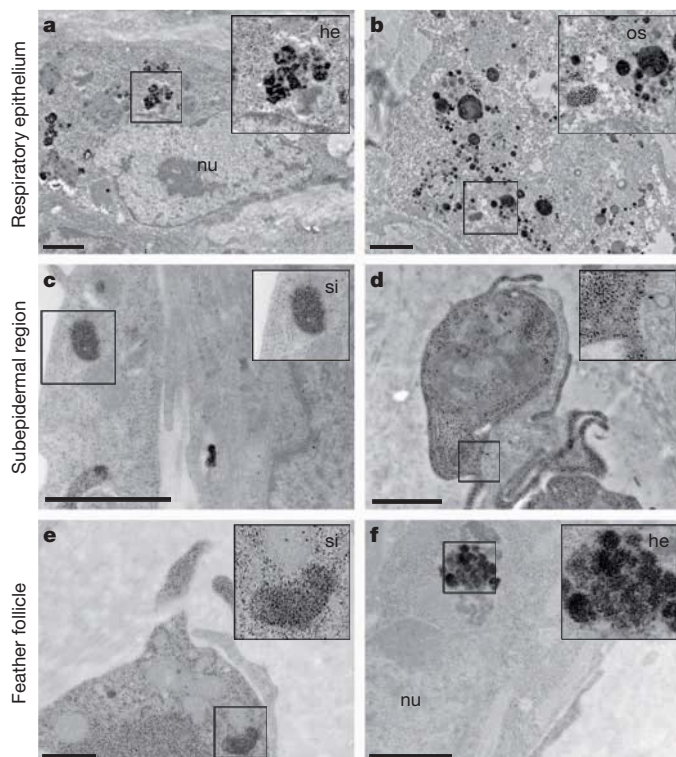


Figure 3 | Ultrastructure of PB-positive cells. a–f, Representative electron micrographs of two different PB-positive cells in the respiratory epithelium (a, b), subepidermal region (c, d) and the feather follicle (e, f; $n = 3$ birds). Cells in all regions were found to contain ferritin-like granules (6–9 nm in diameter) that are present throughout the cytoplasm. In addition we observed haemosiderin (he) clumps and/or membrane bound siderosomes (si). Osmophilic lipid droplets (os) are visible in cells in the respiratory epithelium in b. Cells in the subepidermal region and feather follicle were noted for their slender cytoplasmic projections resembling filopodia, that are seen to engulf a cell in d. Cells in all regions were nucleated (nu). See also Supplementary Fig. 13. Scale bars, 1 μ m.

Fig. 14). Selected area electron diffraction (SAED) failed to identify any cellular structures that contained magnetite, but showed that haemosiderin masses in the feather follicle consist of a goethite-like material ($n = 3$ birds), whereas siderosomes in the respiratory epithelium are comprised of ferrihydrite ($n = 2$ birds) (Supplementary Fig. 15 and Supplementary Table 3). On a cellular level, PB-positive cells in the respiratory epithelium are characterized by the presence of osmophilic lipid vacuoles (Fig. 3b), whereas those cells originating from the feather follicle and subepidermal region had notable dendritic extensions that resembled filopodia (Supplementary Fig. 13d–i, k). In some instances these cytoplasmic tentacles appeared to engulf neighbouring cells, suggesting to us that the PB cells may be phagocytic macrophages (Fig. 3d).

Macrophages are known to reside in the spleen, dermis and respiratory mucosa of multiple species, and to have a vital role in host defence and iron homeostasis²³. Iron accumulates within macrophages during the catabolism of haemoglobin and is stored as ferritin²⁴. In one class of macrophages known as siderophages, ferritin accumulates in membrane-bound siderosomes, which can then undergo proteolytic processing forming haemosiderin. This accumulation of iron renders these cells PB positive^{25,26}. To ascertain whether the PB-positive cells in the upper beak of the pigeon were siderophages, we stained cryosections with sera against major histocompatibility complex class II (MHC II), which labels antigen-presenting cells including macrophages, alongside positive and negative controls ($n \geq 4$ birds) (Fig. 4a–c and Supplementary Fig. 11)²⁷. We observed MHC II co-localization with 98.8% of PB-positive cells in the respiratory epithelium ($n = 104$ cells),

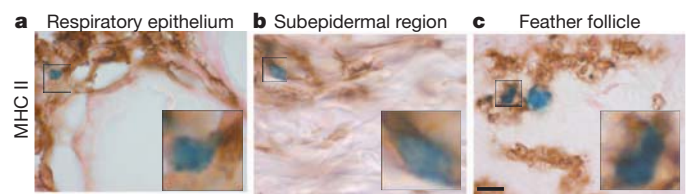


Figure 4 | MHC II immunohistochemistry on PB-positive cells.

a–c, Representative images of coronal sections triple stained with PB, NFR, and sera against the antigen-presenting marker MHC II ($n \geq 4$ birds). MHC II staining was predominant on the surface of cells and found to co-localize with $\geq 94\%$ of PB-positive cells in all regions. Controls are shown in Supplementary Fig. 11. Scale bar, 10 μ m.

95% of PB-positive cells in the subepidermis ($n = 92$ cells), and 94.4% in the feather follicle ($n = 205$ cells). Taken together with our anatomical mapping, subcellular data and neuronal staining, we conclude that clusters of PB-positive cells in the upper beak of the pigeon *C. livia* are macrophages, not magnetosensitive neurons.

As macrophages are not unique to the upper beak, our finding predicts that PB-positive cells should be found throughout the pigeon. To test this we stained skin samples from the back, abdomen, neck, scalp, wing and lower beak of the pigeon ($n \geq 3$ birds). This revealed widespread PB-positive staining in the subepidermis and feather follicle, which was indistinguishable from that observed in the upper beak (Supplementary Fig. 16). Our conclusion further predicts the infiltration of PB-positive macrophages in response to tissue damage or host invasion. We observed such a response in the beak of one of our pigeons (P199), where a large inflammatory lesion with a necrotic centre was surrounded by lymphoplasmacytic cells, including $\sim 80,000$ PB-positive cells that were again characterized by constellations of blue spherules and/or by light blue cytoplasmic staining (Supplementary Fig. 17).

Although we cannot exclude the possibility that a small number of sparsely distributed magnetoreceptors reside at an unknown location in the upper beak of pigeons, this study finds no evidence to support the existence of a subepidermal magnetic sense system that consists of iron-containing dendrites at six specific bilateral loci. This conclusion, which is supported by a critical analysis of the elemental composition of PB-positive cells in the subepidermis²⁸, has several important implications. First, it requires a re-evaluation of behavioural studies that have purported to impair the function of a magnetite-based receptor in the subepidermis of the upper beak and the conclusions that these studies reached^{9,20}. Second, it necessitates a re-assessment as to whether superparamagnetic magnetite has the necessary physical and magnetic properties to act as a magnetosensor in a living system^{10,11,19}. Third, our work reveals that the sensory cells that are responsible for trigeminally mediated magnetic sensation in birds remain undiscovered. These enigmatic cells may reside in the olfactory epithelium, a sensory structure that has been implicated in magnetoreception in the rainbow trout²⁹.

METHODS SUMMARY

Histological studies. We perfused adult pigeons with 4% phosphate-buffered paraformaldehyde (PFA, pH 7.4), and dissected the tissue with ceramic-coated tools. We embedded the tissue in paraffin and prepared 10- μ m coronal sections. We stained sections for 20 min in 5% potassium hexacyanoferrate with 10% HCl, followed by a series of washes and a 2 min exposure to NFR. For immunohistochemistry we incubated the sections with the primary antibody for 18 h in 0.1% Triton PBS with 2–4% milk (pH 7.4), before detection using standard methods, followed by PB staining.

Ultrastructure studies. Adult pigeons were perfused with 2.5% glutaraldehyde supplemented with 2% PFA in PBS (pH 7.4). Tissue was dissected with ceramic-coated tools, incubated for 1 h in phosphate-buffered 2% osmium (pH 7.4), dehydrated and embedded in epoxy resin. We prepared alternative semi-thin (2 μ m) and ultra-thin (70 nm) sections. Semi-thin sections were stained with PB, and the ultra-thin sections were used for TEM, EFTEM and SAED.

Imaging studies. Adult pigeons were perfused with 4% PFA. Following post-fixation and mounting, MRI imaging was performed on a horizontal bore 9.4 T DirectDrive VNMRs system (Agilent Technologies), and CT on a Nucline Nano SPECT/CT imaging system (Mediso).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 December 2011; accepted 19 March 2012.

Published online 11 April 2012.

- Mouritsen, H. & Ritz, T. Magnetoreception and its use in bird navigation. *Curr. Opin. Neurobiol.* **15**, 406–414 (2005).
- Johnsen, S. & Lohmann, K. J. The physics and neurobiology of magnetoreception. *Nature Rev. Neurosci.* **6**, 703–712 (2005).
- Mora, C. V., Davison, M., Wild, J. M. & Walker, M. M. Magnetoreception and its trigeminal mediation in the homing pigeon. *Nature* **432**, 508–511 (2004).
- Kirschvink, J. L., Walker, M. M. & Diebel, C. E. Magnetite-based magnetoreception. *Curr. Opin. Neurobiol.* **11**, 462–467 (2001).
- Fleissner, G. *et al.* Ultrastructural analysis of a putative magnetoreceptor in the beak of homing pigeons. *J. Comp. Neurol.* **458**, 350–360 (2003).
- Falkenberg, G. *et al.* Avian magnetoreception: elaborate iron mineral containing dendrites in the upper beak seem to be a common feature of birds. *PLoS ONE* **5**, e9231 (2010).
- Fleissner, G., Stahl, B., Thalau, P., Falkenberg, G. & Fleissner, G. A novel concept of Fe-mineral-based magnetoreception: histological and physicochemical data from the upper beak of homing pigeons. *Naturwissenschaften* **94**, 631–642 (2007).
- Hanzlik, M. *et al.* Superparamagnetic magnetite in the upper beak tissue of homing pigeons. *Biometals* **13**, 325–331 (2000).
- Wiltchko, R., Schiffner, I., Fuhrmann, P. & Wiltchko, W. The role of the magnetite-based receptors in the beak in pigeon homing. *Curr. Biol.* **20**, 1534–1538 (2010).
- Solov'yov, I. A. & Greiner, W. Theoretical analysis of an iron mineral-based magnetoreceptor model in birds. *Biophys. J.* **93**, 1493–1509 (2007).
- Davila, A. F., Winklhofer, M., Shcherbakov, V. P. & Petersen, N. Magnetic pulse affects a putative magnetoreceptor mechanism. *Biophys. J.* **89**, 56–63 (2005).
- Zapka, M. *et al.* Visual but not trigeminal mediation of magnetic compass information in a migratory bird. *Nature* **461**, 1274–1277 (2009).
- Ritz, T., Thalau, P., Phillips, J. B., Wiltchko, R. & Wiltchko, W. Resonance effects indicate a radical-pair mechanism for avian magnetic compass. *Nature* **429**, 177–180 (2004).
- Cadiou, H. & McNaughton, P. A. Avian magnetite-based magnetoreception: a physiologist's perspective. *J. R. Soc. Interface* **7** (suppl. 2), S193–S205 (2010).
- Beason, R. C. & Semm, P. Magnetic responses of the trigeminal nerve system of the bobolink (*Dolichonyx oryzivorus*). *Neurosci. Lett.* **80**, 229–234 (1987).
- Heyers, D., Zapka, M., Hoffmeister, M., Wild, J. M. & Mouritsen, H. Magnetic field changes activate the trigeminal brainstem complex in a migratory bird. *Proc. Natl Acad. Sci. USA* **107**, 9394–9399 (2010).
- Williams, M. N. & Wild, J. M. Trigeminally innervated iron-containing structures in the beak of homing pigeons, and other birds. *Brain Res.* **889**, 243–246 (2001).
- Wiltchko, W., Munro, U., Ford, H. & Wiltchko, R. Avian orientation: the pulse effect is mediated by the magnetite receptors in the upper beak. *Proc. Biol. Sci.* **276**, 2227–2232 (2009).
- Solov'yov, I. A. & Greiner, W. Micromagnetic insight into a magnetoreceptor in birds: existence of magnetic field amplifiers in the beak. *Phys. Rev. E* **80**, 041919 (2009).
- Stapput, K., Thalau, P., Wiltchko, R. & Wiltchko, W. Orientation of birds in total darkness. *Curr. Biol.* **18**, 602–606 (2008).
- Iancu, T. C. Ferritin and hemosiderin in pathological tissues. *Electron Microsc. Rev.* **5**, 209–229 (1992).
- Richter, G. W. The iron-loaded cell—the cytopathology of iron storage. A review. *Am. J. Pathol.* **91**, 362–404 (1978).
- Wang, J. & Pantopoulos, K. Regulation of cellular iron metabolism. *Biochem. J.* **434**, 365–381 (2011).
- Mebius, R. E. & Kraal, G. Structure and function of the spleen. *Nature Rev. Immunol.* **5**, 606–616 (2005).
- Meguro, R. *et al.* The presence of ferric and ferrous iron in the nonheme iron store of resident macrophages in different tissues and organs: histochemical demonstrations by the perfusion-Perls and -Turnbull methods in the rat. *Arch. Histol. Cytol.* **68**, 171–183 (2005).
- Simson, J. V. & Spicer, S. S. Ferritin particles in macrophages and in associated mast cells. *J. Cell Biol.* **52**, 536–541 (1972).
- Igyártó, B. Z., Lacko, E., Olah, I. & Magyar, A. Characterization of chicken epidermal dendritic cells. *Immunology* **119**, 278–288 (2006).
- Winklhofer, M. & Kirschvink, J. Does avian magnetoreception rely on both magnetite and maghemite? <http://arxiv.org/abs/0805.2249> (2008).
- Walker, M. M. *et al.* Structure and function of the vertebrate magnetic sense. *Nature* **390**, 371–376 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank M. Busslinger, M. Wild and J. Flint for their critical comments on earlier drafts of this manuscript. Thanks also to T. Iancu who commented on our electron micrographs and S. Soto who remarked on the inflammatory lesion in P199. Gratitude is owed to the bio-optics and electron microscopy facilities at the Institute of Molecular Pathology for their assistance in performing experiments. We wish to acknowledge the Centre for Microscopy, Characterisation and Analysis and the Australian Microscopy and Microanalysis Research Facility at the University of Western Australia, a facility funded by the University, State and Commonwealth Governments. Finally, we wish to thank Boehringer Ingelheim, which funds basic science at the Institute of Molecular Pathology.

Author Contributions D.A.K. and C.D.T. conceived and designed the study. M.C.S., C.D.T., M.B., P.P., C.S. and N.E., performed the sectioning, PB staining and counting. D.A.K., C.D.T. and H.C. analysed the resultant data. J.R. and M.L. performed the MRI and CT studies, producing the three-dimensional structure of the pigeon beak. D.A.K. performed the immunohistochemical studies. C.D.T. performed the ultrastructure experiments and J.S. and M.S. did the EFTEM and SAED studies and analysed the data. D.A.K. wrote the paper, and all authors commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.A.K. (keys@imp.ac.at).

METHODS

PB staining. We optimized our PB-staining protocol on our large cohort of pigeons ($n = 172$) that we sourced from seven different lofts. We experimented with a variety of different fixatives, section thickness, and blades before adopting the following protocol. Adult homing pigeons ($n = 12$ Nuremberg cohort, $n = 6$ Vienna cohort) were killed before perfusion with 4% phosphate-buffered paraformaldehyde (PFA, pH 7.4). The tissue was dissected with ceramic-coated tools and post-fixed for 18 h before dehydration in an increasing alcohol series. The tissue was then embedded in paraffin and sectioned coronally (10 μm) with ceramic-coated microtome blades (DuraEdgeHigh Profile, BLM00103P). Four consecutive coronal sections were mounted on positively charged microscope slides (Menzel Superfrost PLUS, Thermo Scientific). For PB staining we incubated every third slide in a freshly prepared solution of 5% potassium hexacyanoferrate (Sigma, P9387) in 10% HCl for 20 min. After three washes in double distilled H_2O , sections were counterstained for 2 min in NFR (Sigma, 60700). Each slide was then scanned (MIRAX Slide Scanner) and PB-positive cells counted manually, using light microscopy where necessary. For anatomical mapping, landmarks were identified (Supplementary Fig. 5), and 500- μm -thick increments determined. To obtain estimated total cell counts the number of PB-positive cells counted within a normalized increment were divided by the number of sections counted within that increment and multiplied by a factor of 50 (see Supplementary Tables 1 and 2). We compared the number of PB-positive cells in males and females by performing a Student's *t*-test. All pigeons were sexed using genetic methods as previously described³⁰, and experiments were performed in accordance with the relevant guidelines and regulations (Magistrat 60, Veterinäramt, MA60-001603/2010/002).

Immunohistochemistry. For staining with neuronal antibodies slides were deparaffinized and washed in PBS (pH 7.4) before incubation with the primary antibody for 18 h in 0.1% Triton PBS with 2% milk (pH 7.4). Primary antibodies were used at the following concentrations: NF (Millipore, MAB1621, 1:2,000), TUBB3 (Covance, MMS-435P, 1:1,000), MAP1B (Santa Cruz, SC-58784, 1:75). Following a series of washes in PBS, slides were incubated for 2 h with a biotinylated secondary antibody (1:500), before visualization with a peroxidase-based Vectastain Elite ABC kit (Vector Labs, PK-4002) and the chromophore DAB (3,5-diaminobenzidine, Dako). For MHCII staining, 10- μm cryosections were prepared, quenched in 2% H_2O_2 in PBS for 30 min, before incubation for 18 h in 0.1% Triton PBS with 4% milk (pH 7.4) with the primary antibody (Santa Cruz, SC-59323, 1:500). This antibody is a mouse monoclonal antibody raised against white blood cells originating from the chicken. To avoid cross reaction with endogenous biotin/avidin, a HRP-conjugated secondary antibody was used (Biorad, 1:500), and staining visualized with DAB. Sections were thoroughly washed with PBS before PB staining and scanning as described above. All cell counting and co-localization studies were performed blinded to the antibody used. The overall percentage of co-localization was determined by calculating the rate of co-localization per bird, and ascertaining the mean.

Ultrastructure studies. Adult pigeons were perfused with 2.5% glutaraldehyde supplemented with 2% PFA in PBS (pH 7.4) (Glut-PFA), before tissue dissection with ceramic-coated tools. Following post-fixation for 48 h, this tissue was washed with PBS, incubated for 1 h in phosphate-buffered 2% osmium (pH 7.4), dehydrated and embedded in epoxy resin. After polymerization, the blocks were trimmed and sectioned, alternatively taking semi-thin (2 μm) and ultra-thin (70 nm) sections. The ultra-thin sections were mounted on formvar-film copper slot grids for TEM, whereas the semi-thin sections were etched with 21% sodium ethoxide in ethanol, rehydrated and stained with PB and NFR. Where necessary, PB staining was intensified by incubating the sections in 0.5% DAB³¹. TEM imaging on ultra-thin sections used a 100 kV electron microscope (FEI Morgagni 268D) with a CCD camera (Morada Olympus-SIS). For EFTEM imaging and selected area electron diffraction, the ultra-thin sections were mounted on holey carbon-film copper finder grids (Quantifoil, R3.5/1) and analysed on a 200 kV TEM (JEOL, 2100) fitted with a Gatan Imaging Filter (Tridiem) and CCD camera (Orion SC1000). For EFTEM, bright-field images were taken before obtaining elemental maps for iron, which were acquired using the iron M-edge and generated using the conventional three-window method. Two pre-edge (background) images were acquired at energies of 45 and 50 eV, and the post-edge (signal energy) image was acquired by centring the filter's energy-selecting slit at 59 eV with a slit width of 5 eV (~ 10 s acquisition time). Diffraction data was obtained from iron deposits identified by EFTEM and the data calibrated against a polycrystalline gold standard.

MRI. Animals were killed and perfused with 4% PFA as described above. Following 18 h of post-fixation, heads were mounted in a 70-mm diameter PE tube filled with proton-free perfluoro-polyether fomblin (Solvay Solexis S.p.A.). Imaging was performed on a horizontal bore 9.4 T DirectDrive VNMRS system (Agilent Technologies) using a 72 mm quadrature birdcage volume coil (RAPID Biomedical GmbH). For three-dimensional imaging, a gradient-echo sequence with the following parameters was used: time to echo (TE), 2.8 ms; time to repetition (TR), 280 ms; flip angle, 40°; six averages, field of view, 70 \times 70 \times 35 mm; matrix size, 512 \times 512 \times 256. For higher-resolution images of the beak, PFA-fixed beaks were incubated with 8 mM gadolinium solution (Magnevist, Bayer AG) in PBS for 48 h followed by embedding in agar containing 8 mM gadolinium. Gradient-echo images were acquired using the following parameters: TE, 2.7 ms; TR, 25 ms; flip angle, 45°; five averages; field of view, 28 \times 28 \times 35 mm; matrix size, 560 \times 560 \times 700. Regions of interest were segmented using thresholds with manual adjustments where necessary using Amira visualization software (v.5.2.2, Visage Imaging).

CT imaging. CT imaging was performed on a Nucline Nano SPECT/CT imaging system (Mediso) using the following imaging parameters: 360 projection, pitch 0.5, 55 kVp, 145 μA acquired at 45 μm isotropic reconstructed to 50 μm isotropic.

30. Horng, Y. M., Wu, C. P., Wang, Y. C. & Huang, M. C. A novel molecular genetic marker for gender determination of pigeons. *Theriogenology* **65**, 1759–1768 (2006).
31. Moos, T. & Mollgard, K. A sensitive post-DAB enhancement technique for demonstration of iron in the central nervous system. *Histochemistry* **99**, 471–475 (1993).

Spatial partitioning of the regulatory landscape of the X-inactivation centre

Elphège P. Nora^{1,2,3}, Bryan R. Lajoie^{4*}, Edda G. Schulz^{1,2,3*}, Luca Giorgetti^{1,2,3*}, Ikuhiro Okamoto^{1,2,3}, Nicolas Servant^{1,5,6}, Tristan Piolot^{1,2,3}, Nynke L. van Berkum⁴, Johannes Meisig⁷, John Sedat⁸, Joost Gribnau⁹, Emmanuel Barillot^{1,5,6}, Nils Blüthgen⁷, Job Dekker⁴ & Edith Heard^{1,2,3}

In eukaryotes transcriptional regulation often involves multiple long-range elements and is influenced by the genomic environment¹. A prime example of this concerns the mouse X-inactivation centre (*Xic*), which orchestrates the initiation of X-chromosome inactivation (XCI) by controlling the expression of the non-protein-coding *Xist* transcript. The extent of *Xic* sequences required for the proper regulation of *Xist* remains unknown. Here we use chromosome conformation capture carbon-copy (5C)² and super-resolution microscopy to analyse the spatial organization of a 4.5-megabases (Mb) region including *Xist*. We discover a series of discrete 200-kilobase to 1 Mb topologically associating domains (TADs), present both before and after cell differentiation and on the active and inactive X. TADs align with, but do not rely on, several domain-wide features of the epigenome, such as H3K27me3 or H3K9me2 blocks and lamina-associated domains. TADs also align with coordinately regulated gene clusters. Disruption of a TAD boundary causes ectopic chromosomal contacts and long-range transcriptional misregulation. The *Xist/Tsix* sense/antisense unit illustrates how TADs enable the spatial segregation of oppositely regulated chromosomal neighbourhoods, with the respective promoters of *Xist* and *Tsix* lying in adjacent TADs, each containing their known positive regulators. We identify a novel distal regulatory region of *Tsix* within its TAD, which produces a long intervening RNA, *Linx*. In addition to uncovering a new principle of *cis*-regulatory architecture of mammalian chromosomes, our study sets the stage for the full genetic dissection of the X-inactivation centre.

The X-inactivation centre was originally defined by deletions and translocations as a region spanning several megabases^{3,4}, and contains several elements known to affect *Xist* activity, including its repressive antisense transcript *Tsix* and its regulators *Xite*, *DXPas34* and *Tsx*^{5,6}. However, additional control elements must exist, as single-copy transgenes encompassing *Xist* and up to 460 kb of flanking sequences are unable to recapitulate proper *Xist* regulation⁷. To characterize the *cis*-regulatory landscape of the *Xic* in an unbiased approach, we performed 5C² across a 4.5-Mb region containing *Xist*. We designed 5C-Forward and 5C-Reverse oligonucleotides following an alternating scheme², thereby simultaneously interrogating nearly 250,000 possible chromosomal contacts in parallel, with a mean resolution of 10–20 kb (Fig. 1a; see Supplementary Methods). Analysis of undifferentiated mouse embryonic stem cells (ESCs) revealed that long-range (>50 kb) contacts preferentially occur within a series of discrete genomic blocks, each covering 0.2–1 Mb (Fig. 1b). These blocks differ from the higher-order organization recently observed by Hi-C⁸, corresponding to much larger domains of open or closed chromatin, that come together in the nucleus to form A and B types of compartments⁸. Instead, our

5C analysis shows self-associating chromosomal domains occurring at the sub-megabase scale. The size and location of these domains is identical in male and female mouse ESCs (Supplementary Fig. 1) and in different mouse ESC lines (Supplementary Fig. 2 and Supplementary Data 1).

To examine this organization with an alternative approach, we performed three-dimensional DNA fluorescent *in situ* hybridization (FISH) in male mouse ESCs. Nuclear distances were found to be significantly shorter between probes lying in the same 5C domain than in different domains (Fig. 1c, d), and a strong correlation was found between three-dimensional distances and 5C counts (Supplementary Fig. 3a, b). Furthermore, using pools of tiled bacterial artificial chromosome (BAC) probes spanning up to 1 Mb and structured illumination microscopy, we found that large DNA segments belonging to the same 5C domain colocalize to a greater extent than DNA segments located in adjacent domains (Fig. 1e), and this throughout the cell cycle (Supplementary Fig. 3c, d). Based on 5C and FISH data, we conclude that chromatin folding at the sub-megabase scale is not random, and partitions this chromosomal region into a succession of topologically associating domains (TADs).

We next investigated what might drive chromatin folding in TADs. We first noticed a striking alignment between TADs and the large blocks of H3K27me3 and H3K9me2 (ref. 9) that are known to exist throughout the mammalian genomes^{10–13} (for example, TAD E, Fig. 2 and Supplementary Fig. 4). We therefore examined 5C profiles of *G9a*^{-/-} (also known as *Ehmt2*) mouse ESCs, which lack H3K9me2, notably at the *Xic*¹⁴, and *Eed*^{-/-} mouse ESCs, which lack H3K27me3 (ref. 15). No obvious change in overall chromatin conformation was observed, and TADs were not affected either in size or position in these mutants (Fig. 2 and Supplementary Fig. 4b). Thus TAD formation is not due to domain-wide H3K27me3 or H3K9me2 enrichment. Instead, such segmental chromatin blocks might actually be delimited by the spatial partitioning of chromosomes into TADs.

We then addressed whether folding in TADs is driven by discrete boundary elements at their borders. 5C was performed in a mouse ESC line carrying a 58-kb deletion (Δ TX¹⁶), encompassing the boundary between the *Xist* and *Tsix* TADs (D and E; Fig. 2b). We observed ectopic contacts between sequences in TADs D and E and an altered organization of TAD E. Boundary elements can thus mediate the spatial segregation of neighbouring chromosomal segments. Within the TAD D–E boundary, a CTCF-binding site was recently implicated in insulating *Tsix* from remote regulatory influences¹⁷. However, alignment of CTCF- and cohesin-binding sites in mouse ESCs¹⁸ with our 5C data showed that, although these factors are present at most TAD boundaries (Supplementary Fig. 4), they are also frequently present within TADs, excluding them as the sole determinants of TAD

¹Institut Curie, 26 rue d'Ulm, Paris F-75248, France. ²CNRS UMR3215, Paris F-75248, France. ³INSERM U934, Paris F-75248, France. ⁴Programs in Systems Biology and Gene Function and Expression, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605-0103, USA. ⁵INSERM U900, Paris, F-75248 France. ⁶Mines ParisTech, Fontainebleau, F-77300 France. ⁷Institute of Pathology, Charité-Universitätsmedizin, 10117 Berlin, and Institute of Theoretical Biology Humboldt Universität, 10115 Berlin, Germany.

⁸Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California 94158-2517, USA. ⁹Department of Reproduction and Development, Erasmus MC, University Medical Center, 3000 CA Rotterdam, The Netherlands.

*These authors contributed equally to this work.

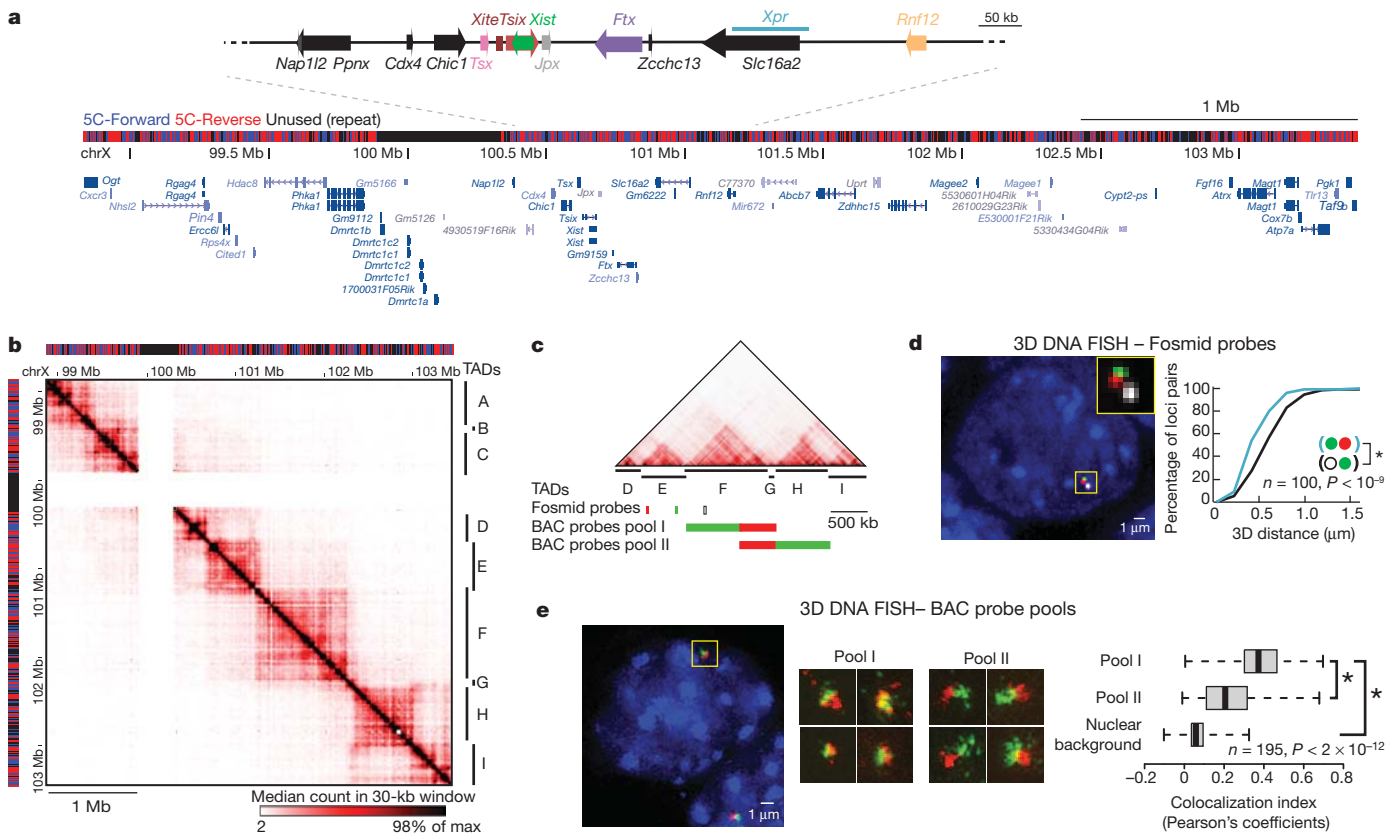


Figure 1 | Chromosome partitioning into topologically associating domains (TADs). **a**, Distribution of 5C-Forward and 5C-Reverse HindIII restriction fragments across the 4.5 Mb analysed showing positions of RefSeq genes and known XCI regulatory loci. **b**, 5C data sets from XY undifferentiated mouse ESCs (E14), displaying median counts in 30-kb windows every 6 kb. Chromosomal contacts are organized into discrete genomic blocks (TADs A–I). A region containing segmental duplications excluded from the 5C analysis is masked (white). **c**, Positions of DNA FISH probes. **d**, Interphase

positioning. Furthermore, the fact that the two neighbouring domains do not merge completely in Δ TX cells (Fig. 2b) implies that additional elements, within TADs, can act as relays when a main boundary is removed. The factors underlying an element's capacity to act as a canonical or shadow boundary remain to be investigated.

Next we asked whether TAD organization changes during differentiation or XCI. Both male neuronal progenitors cells (NPCs) and male primary mouse embryonic fibroblasts (MEFs) show similar organization to mouse ESCs, with no obvious change in TAD positioning. However, consistent differences in the internal contacts within TADs were observed (Fig. 3a, Supplementary Figs 2 and 5). Noticeably, some TADs were found to become lamina-associated domains¹⁹ (LADs) at certain developmental stages (Fig. 3b). Thus chromosome segmentation into TADs reveals a modular framework where changes in chromatin structure or nuclear positioning can occur in a domain-wide fashion during development.

We then assessed TAD organization on the inactive X, by combining *Xist* RNA FISH, to identify the inactive X, and super-resolution DNA FISH using BAC probe pools on female MEFs. We found that colocalization indices on the inactive X were still higher for sequences belonging to the same TAD than for neighbouring TADs (Supplementary Fig. 6a). However, the difference was significantly lower for the inactive X than for the active X. Deconvolution of the respective contributions of the active X and inactive X in 5C data from female MEFs (see Supplementary Methods and Supplementary Fig. 6) similarly revealed that global organization in TADs remains on the inactive X, albeit in a much attenuated form, but that specific long-range

nuclear distances are smaller for probes in the same 5C domain. **e**, Structured illumination microscopy reveals that colocalization of neighbouring sequences is greater when they belong to the same 5C domain. Boxplots show the distribution of Pearson's correlation coefficient between red and green channels, with whiskers and boxes encompassing all and 50% of values, respectively; central bars denote the median correlation coefficient. Statistical significance was assessed using Wilcoxon's rank sum test.

contacts within TADs are lost. This, together with a recent report focused on longer-range interactions²⁰, suggests that the inactive X has a more random chromosomal organization than its active homologue, even below the megabase scale.

We next investigated how TAD organization relates to gene expression dynamics during early differentiation. A transcriptome analysis, consisting of microarray measurements at 17 time points over the first 84 h of female mouse ESC differentiation was performed (Fig. 4a). During this time window, most genes in the 5C region were either up- or downregulated. Statistical analysis demonstrated that expression profiles of genes with promoters located within the same TAD are correlated (Fig. 4b). This correlation (median correlation coefficient cc of 0.40) is significantly higher than for genes in different domains (cc of 0.03, $P < 10^{-9}$) or for genes across the X chromosome in randomly selected, TAD-size regions (cc of 0.09, $P < 10^{-7}$). The observed correlations within TADs seem not to depend on distance between genes, and are thus distinct from previously described correlations between neighbouring genes²¹ that decay on a length scale of approximately 100 kb (Supplementary Fig. 7). Our findings indicate that physical clustering within TADs may be used to coordinate gene expression patterns during development. Furthermore, deletion of the boundary between *Xist* and *Tsix* in Δ XTX cells was accompanied by long-range transcriptional misregulation (Supplementary Fig. 8), underlining the role that chromosome partitioning into TADs can play in long-range transcriptional control.

A more detailed analysis of each domain (Supplementary Fig. 7) revealed that co-expression is particularly pronounced in TADs D, E

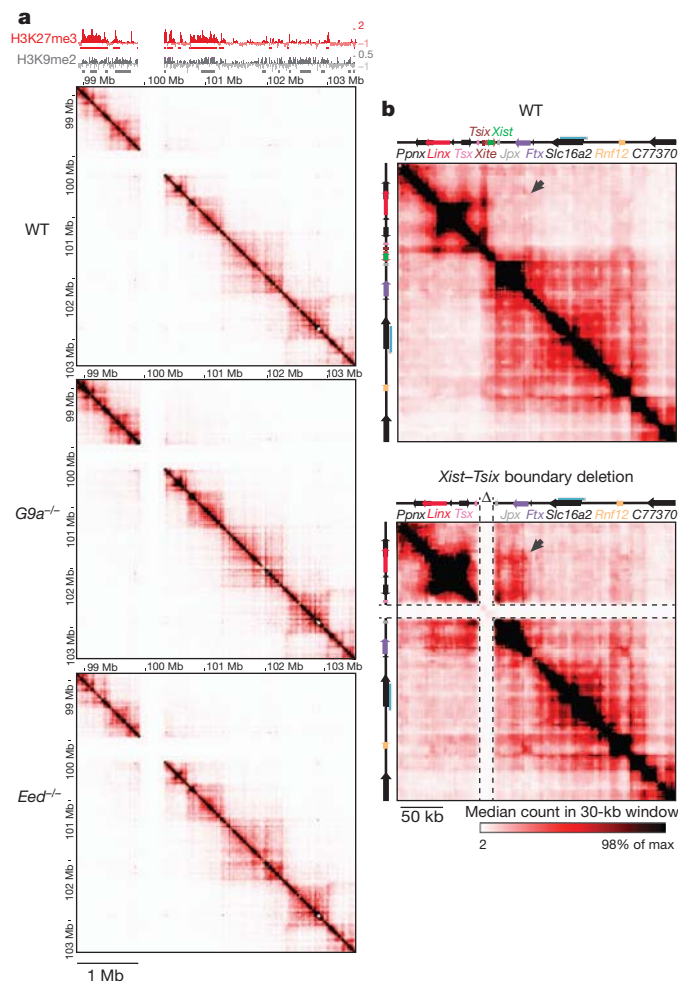


Figure 2 | Determinants of topologically associating domains. **a**, Blocks of contiguous enrichment in H3K27me3 or H3K9me2 (ref. 11) align with the position of TADs (chromatin immunoprecipitation on chip from ref. 9) in wild-type cells (TT2), but TADs are largely unaffected in the absence of H3K9me2 in male *G9a*^{-/-} cells or H3K27me3 in male *Eed*^{-/-} cells. **b**, Deletion of a boundary at *Xist/Tsix* disrupts folding pattern of the two neighbouring TADs.

and F (Fig. 4b, c). Although correlations are strongest within TADs, there is some correlation between TADs showing the same trend, such as TADs D and F, which are both downregulated during differentiation. Only TAD E, which contains *Xist* and all of its known positive

regulators *Jpx*, *Ftx*, *Xpr/Xpct* and *Rnf12*⁵ (*Jpx*, *Ftx*, *Xpct* and *Rnf12* are also known as *Enox*, *B230206F22Rik*, *Slc16a2* and *Rlim*, respectively) is anti-correlated with most other genes in the 4.5 Mb region, being upregulated during differentiation (Supplementary Fig. 7). The fact that these coordinately upregulated loci are located in the same TAD suggests that they are integrated into a similar *cis*-regulatory network, potentially sharing common *cis*-regulatory elements. We therefore predict that TAD E (~550 kb) represents the minimum 5' regulatory region required for accurate *Xist* expression, explaining why even the largest transgenes tested so far (covering 150 kb 5' to *Xist*, Fig. 5a) cannot recapitulate normal *Xist* expression⁷.

The respective promoters of *Xist* and *Tsix* lie in two neighbouring TADs with transcription crossing the intervening boundary (Fig. 2b), consistent with previous 3C experiments²². Whereas the *Xist* promoter and its positive regulators are located in TAD E, the promoter of its antisense repressor, *Tsix*, lies in TAD D, which extends up to *Ppnlx* (also known as *4930519F16Rik*)/*Nap1l2*, more than 200 kb away (Fig. 2b). Thus, in addition to the *Xite* enhancer, more distant elements within TAD D may participate in *Tsix* regulation. To test this we used two different single-copy transgenic mouse lines, Tg53 and Tg80 (ref. 23). Both transgenes contain *Xist*, *Tsix* and *Xite* (Fig. 5a). Tg53 encompasses the whole of TAD D, whereas Tg80 is truncated just 5' to *Xite* (Fig. 5a and Supplementary Fig. 9). In the inner cell mass of male mouse embryos at embryonic day 4.0 (E4.0), *Tsix* transcripts could be readily detected from Tg53, as well as from the endogenous X (Fig. 5b). However, no *Tsix* expression could be detected from Tg80, which lacks the distal portion of TAD D (Fig. 5b). Thus, sequences within TAD D must contain essential elements for the correct developmental regulation of *Tsix*.

Within TAD D, several significant looping events involving the *Tsix* promoter or its enhancer *Xite* were detected (Figs 2b and 5a, Supplementary Fig. 10). Alignment of 5C maps with chromatin signatures of enhancers in mouse ESCs (Supplementary Fig. 11) suggested the existence of multiple regulatory elements within this region. We also identified a transcript initiating approximately 50 kb upstream of the *Ppnlx* promoter (Fig. 5a), from a region bound by pluripotency factors and corresponding to a predicted promoter for a large (80 kb) intervening non-coding RNA (lincRNA²⁴, Supplementary Fig. 12) which we termed *Linx* (large intervening transcript in the *Xic*). *Linx* RNA shares several features with non-coding RNAs, such as accumulation around its transcription site²⁵ (Fig. 5c), nuclear enrichment and abundance of the unspliced form²⁶ (Supplementary Fig. 12 and 13). *Linx* and *Tsix* are co-expressed in the inner cell mass of blastocysts from E3.5–4.0 onwards, as well as in male and female mouse ESCs (Fig. 5c). *Linx* RNA is not detected earlier in embryogenesis, nor in extra-embryonic lineages, implying an epiblast-specific function

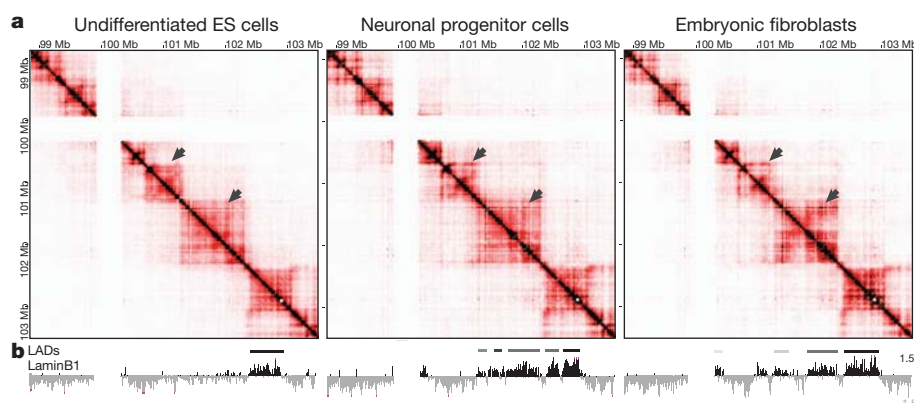


Figure 3 | Dynamics of topologically associating domains during cell differentiation. **a**, Comparison of 5C data from male mouse ESCs (E14), NPCs (E14) and primary MEFs reveals general conservation of TAD positions during differentiation, but differences in their internal organization (arrows highlight

examples of tissue-specific patterns). **b**, Lamina-associated domains (LADs, from ref. 19) align with TADs. Chromosomal positions of tissue-specific LADs reflect gain of lamina association by TADs, as well as internal reorganization of lamina-associated TADs during differentiation.

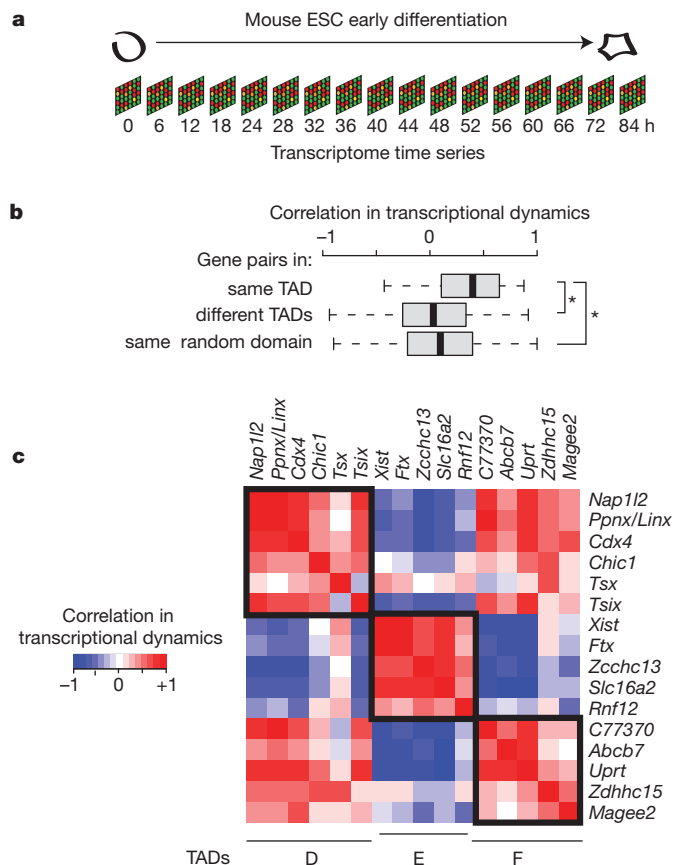


Figure 4 | Transcriptional co-regulation within topologically associating domains. **a**, Female mouse ESCs were differentiated towards the epiblast stem cell lineage for 84 h. Transcript levels were measured every 4–6 h at 17 different time points by microarray analysis. **b**, Pearson's correlation coefficients over all time points were calculated for gene pairs lying in the same TAD, pairs in different TADs and for pairs in randomly defined domains on the X chromosome that contain a similar number of genes and are of comparable size. Boxplots show the distribution of Pearson's correlation coefficients, with whiskers and boxes encompassing all and 50% of values, respectively, and central bars denoting the median correlation coefficient. * represents significant difference with $P < 10^{-7}$ using Wilcoxon's rank sum test. **c**, Pearson's correlation coefficients for gene pairs in TADs D, E and F with red denoting positive and blue negative correlation. Boxes indicate the TAD boundaries.

(Supplementary Fig. 9). Triple RNA FISH for *Linx*, *Tsix* and *Xist* in differentiating female mouse ESCs (Supplementary Fig. 14) revealed that before *Xist* upregulation, the probability of *Tsix* expression from alleles co-expressing *Linx* is significantly higher than from alleles that do not express *Linx* (Fig. 5d). Furthermore, *Linx* expression is frequently monoallelic, even before *Xist* upregulation (Supplementary Fig. 14), revealing a transcriptional asymmetry of the two *Xic* alleles before XCI. Taken together, our experiments based on 5C, transgenesis and RNA FISH, point towards a role for *Linx* in the long-range transcriptional regulation of *Tsix* — either through its chromosomal association with *Xite* and/or via the RNA it produces. This analysis of the *Xist/Tsix* region illustrates how spatial compartmentalization of chromosomal neighbourhoods in TADs partitions the *Xic* into two large regulatory domains, with opposite transcriptional fates (Supplementary Fig. 15).

In conclusion, our study reveals that sub-megabase folding of mammalian chromosomes results in the self-association of large chromosomal neighbourhoods in the three-dimensional space of the nucleus. The stability of such partitioning throughout differentiation, X inactivation and in cell lines with impaired histone-modifying machineries, indicates that this level of chromosomal organization may provide a basic framework onto which other domain-wide

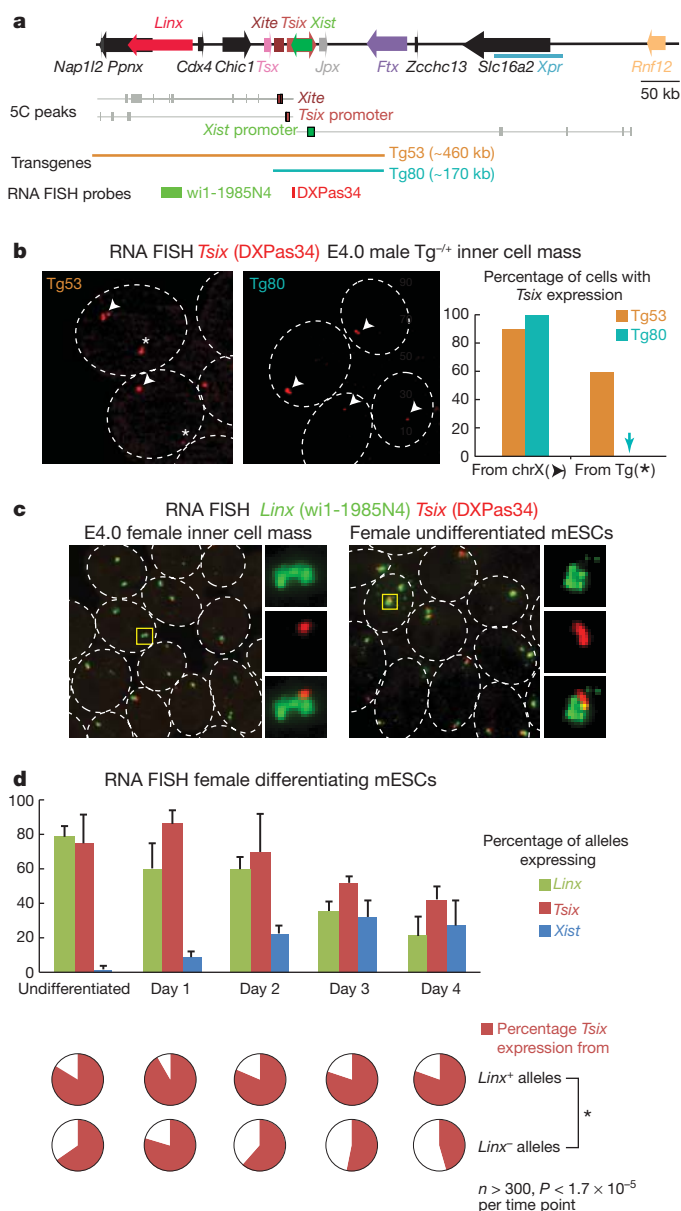


Figure 5 | 5C maps reveal new regulatory regions in the *Xic*. **a**, Statistically significant looping events (5C peaks) for restriction fragments within *Xite*, *Tsix* promoter or *Xist* promoter within their respective TAD, in male (E14) mouse ESCs. The Tg80 YAC transgene lacks genomic elements found to interact physically with *Xite/Tsix* that are present in Tg53. **b**, RNA FISH analysis of *Tsix* expression is detected in the inner cell masses of heterozygous transgenic male E4.0 embryos by RNA FISH from single-copy paternally inherited Tg53 but not Tg80 transgenes. Transgenic (star) and endogenous *Tsix* alleles (arrowhead) were discriminated by subsequent DNA FISH as in Supplementary Fig. 5. $n = 20$ inner cell mass cells (two embryos each). **c**, *Linx* transcripts (green, wi1-1985N4 probe) are expressed in both E4.0 inner cell mass cells and mouse ESCs, together with *Tsix* (red, DXPas34 probe), and unspliced transcripts accumulate locally in a characteristic cloud-like shape. **d**, RNA FISH in differentiating female mouse ESCs revealing synchronous downregulation of *Linx* and *Tsix* with concomitant upregulation of *Xist* (detected with a strand-specific probe). Bars are the standard deviation around the mean of three experiments. Triple-colour RNA FISH allows simultaneous detection of *Linx*, *Tsix* and *Xist* RNAs. Scoring of *Xist*-negative alleles demonstrates that before *Xist* upregulation *Tsix* expression is more frequent from *Linx*-expressing alleles than from *Linx* non-expressing alleles, at all time points tested. Presented is the mean of three experiments. Statistical differences were assessed using Fisher's exact test. Cells were differentiated in monolayers by withdrawal of leukaemia inhibitory factor (LIF).

features, such as lamina association and blocks of histone modification, can be dynamically overlaid. Our data also point to a role for TADs in shaping regulatory landscapes, by defining the extent of sequences that belong to the same regulatory neighbourhood. We anticipate that TADs may underlie regulatory domains previously proposed on the basis of functional and synteny conservation studies^{27,28}. We believe that the principles we have revealed here will not be restricted to the *Xic*, as spatial partitioning of chromosomal neighbourhoods occurs throughout the genome of mouse and human²⁹, as well as *Drosophila*³⁰ and *E. coli*³¹. We have shown that TAD boundaries can have a critical role in high-order chromatin folding and proper long-range transcriptional control. Future work will clarify the mechanisms driving this level of chromosomal organization, and to what extent it generally contributes to transcriptional regulation. In summary, our study provides new insights into the *cis*-regulatory architecture of chromosomes that orchestrates transcriptional dynamics during development, and paves the way to dissecting the constellation of control elements of *Xist* and its regulators within the *Xic*.

METHODS SUMMARY

5C was performed on mouse ESCs, mouse NPCs and primary MEFs following a previously described protocol² with modifications, and sequenced on one lane of an Illumina GAIIX. RNA and DNA FISH were performed on mouse ESCs and inner cell masses extracted from pre-implantation embryos as previously described⁷, with modifications. Full experimental and bioinformatic methods are detailed in Supplementary Information.

Received 3 October 2011; accepted 22 March 2012.

Published online 11 April 2012.

- Kleinjan, D. A. & Lettice, L. A. Long-range gene control and genetic disease. *Adv. Genet.* **61**, 339–388 (2008).
- Dostie, J. *et al.* Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006).
- Rastan, S. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos—location of the inactivation centre. *J. Embryol. Exp. Morphol.* **78**, 1–22 (1983).
- Rastan, S. & Robertson, E. J. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J. Embryol. Exp. Morphol.* **90**, 379–388 (1985).
- Augui, S., Nora, E. P. & Heard, E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nature Rev. Genet.* **12**, 429–442 (2011).
- Anguera, M. C. *et al.* *Tsx* produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. *PLoS Genet.* **7**, e1002248 (2011).
- Heard, E., Mongelard, F., Arnaud, D. & Avner, P. *Xist* yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. *Mol. Cell. Biol.* **19**, 3156–3156 (1999).
- Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- Marks, H. *et al.* High-resolution analysis of epigenetic changes associated with X inactivation. *Genome Res.* **19**, 1361–1373 (2009).
- Pauler, F. M. *et al.* H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res.* **19**, 221–233 (2009).
- Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature Genet.* **41**, 246–250 (2009).
- Lienert, F. *et al.* Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet.* **7**, e1002090 (2011).
- Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479–491 (2010).
- Rougeulle, C. *et al.* Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol. Cell. Biol.* **24**, 5475–5484 (2004).
- Montgomery, N. D. *et al.* The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr. Biol.* **15**, 942–947 (2005).
- Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F. & Gribnau, J. X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. *Cell* **132**, 410–421 (2008).
- Spencer, R. J. *et al.* A boundary element between *Tsix* and *Xist* binds the chromatin insulator Ctf and contributes to initiation of X chromosome inactivation. *Genetics* CrossRef (2011).
- Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
- Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
- Splinter, E. *et al.* The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on *Xist* RNA. *Genes Dev.* **25**, 1371–1383 (2011).
- Caron, H. *et al.* The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291**, 1289–1292 (2001).
- Tsai, C.-L., Rowntree, R. K., Cohen, D. E. & Lee, J. T. Higher order chromatin structure at the X-inactivation center via looping DNA. *Dev. Biol.* **319**, 416–425 (2008).
- Heard, E. *et al.* Transgenic mice carrying an *Xist*-containing YAC. *Hum. Mol. Genet.* **5**, 441–450 (1996).
- Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
- Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl Acad. Sci. USA* **106**, 11667–11672 (2009).
- Seidl, C. I. M., Stricker, S. H. & Barlow, D. P. The imprinted *Air* ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J.* **25**, 3565–3575 (2006).
- Ruf, S. *et al.* Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nature Genet.* **43**, 379–386 (2011).
- Kikuta, H. *et al.* Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res.* **17**, 545–555 (2007).
- Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* doi:10.1038/nature11082 (this issue).
- Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
- Mercier, R. *et al.* The MatP/matS site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135**, 475–485 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Pollex and T. Forné for experimental help; the imaging facility PICTIBISA@BDD for technical assistance, D. Gentien and C. Hego for microarray hybridizations. We thank K. Bernhard, F. Stewart and A. Smith for protocols and material for 2i culture and EpiSC differentiation. We are grateful to members of the E.H. laboratory for critical input. This work was funded by grants from the Ministère de la Recherche et de l'Enseignement Supérieur and the ARC (to E.P.N.); a HFSP Long term fellowship (LT000597/2010-L) (to E.G.S.); EU EpiGeneSys FP7 Network of Excellence no. 257082, the Fondation pour la Recherche Médicale, ANR, ERC Advanced Investigator award no. 250367 and EU FP7 SYBOSS grant no. 242129 (to E.H.). N.B. was supported by BMBF (FORSSYS) and EMBO (fellowship ASTF 307-2011). J.D., B.R.L. and N.L.v.B. were supported by NIH (R01 HG003143) and a W. M. Keck Foundation Distinguished Young Scholar Award.

Author Contributions E.P.N. performed and analysed 3C, 5C, (RT)-qPCR, immunofluorescence, RNA and DNA FISH. B.R.L. and N.L.v.B. helped in the design and/or the analysis of 3C and 5C. L.G. performed 3C, FISH and 5C analysis. E.G.S. generated the time-course transcriptomic data, which was analysed by J.M. and N.B.; I.O. performed FISH on pre-implantation embryos. J.G. donated the TTX mouse ESC line. N.S. and E.B. helped in the epigenomic and 5C analyses. J.S. and T.P. set up OMX microscopy and analysis and T.P. performed structured illumination microscopy and image analysis. The manuscript was written by E.P.N., J.D. and E.H. with contribution from E.G.S. and input from all authors.

Author Information High-throughput data are deposited in Gene Expression Omnibus under accession number GSE35721 for all 5C experiments and GSE34243 for expression microarrays. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.H. (edith.heard@curie.fr) or J.D. (job.dekker@umassmed.edu).

Topological domains in mammalian genomes identified by analysis of chromatin interactions

Jesse R. Dixon^{1,2,3}, Siddarth Selvaraj^{1,4}, Feng Yue¹, Audrey Kim¹, Yan Li¹, Yin Shen¹, Ming Hu⁵, Jun S. Liu⁵ & Bing Ren^{1,6}

The spatial organization of the genome is intimately linked to its biological function, yet our understanding of higher order genomic structure is coarse, fragmented and incomplete. In the nucleus of eukaryotic cells, interphase chromosomes occupy distinct chromosome territories, and numerous models have been proposed for how chromosomes fold within chromosome territories¹. These models, however, provide only few mechanistic details about the relationship between higher order chromatin structure and genome function. Recent advances in genomic technologies have led to rapid advances in the study of three-dimensional genome organization. In particular, Hi-C has been introduced as a method for identifying higher order chromatin interactions genome wide². Here we investigate the three-dimensional organization of the human and mouse genomes in embryonic stem cells and terminally differentiated cell types at unprecedented resolution. We identify large, megabase-sized local chromatin interaction domains, which we term ‘topological domains’, as a pervasive structural feature of the genome organization. These domains correlate with regions of the genome that constrain the spread of heterochromatin. The domains are stable across different cell types and highly conserved across species, indicating that topological domains are an inherent property of mammalian genomes. Finally, we find that the boundaries of topological domains are enriched for the insulator binding protein CTCF, housekeeping genes, transfer RNAs and short interspersed element (SINE) retrotransposons, indicating that these factors may have a role in establishing the topological domain structure of the genome.

To study chromatin structure in mammalian cells, we determined genome-wide chromatin interaction frequencies by performing the Hi-C experiment² in mouse embryonic stem (ES) cells, human ES cells, and human IMR90 fibroblasts. Together with Hi-C data for the mouse cortex generated in a separate study (Y. Shen *et al.*, manuscript in preparation), we analysed over 1.7-billion read pairs of Hi-C data corresponding to pluripotent and differentiated cells (Supplementary Table 1). We normalized the Hi-C interactions for biases in the data (Supplementary Figs 1 and 2)³. To validate the quality of our Hi-C data, we compared the data with previous chromosome conformation capture (3C), chromosome conformation capture carbon copy (5C), and fluorescence *in situ* hybridization (FISH) results^{4–6}. Our IMR90 Hi-C data show a high degree of similarity when compared to a previously generated 5C data set from lung fibroblasts (Supplementary Fig. 4). In addition, our mouse ES cell Hi-C data correctly recovered a previously described cell-type-specific interaction at the *Phc1* gene⁵ (Supplementary Fig. 5). Furthermore, the Hi-C interaction frequencies in mouse ES cells are well-correlated with the mean spatial distance separating six loci as measured by two-dimensional FISH⁶ (Supplementary Fig. 6), demonstrating that the normalized Hi-C data can accurately reproduce the expected nuclear distance using an independent method. These results demonstrate that our Hi-C data are of

high quality and accurately capture the higher order chromatin structures in mammalian cells.

We next visualized two-dimensional interaction matrices using a variety of bin sizes to identify interaction patterns revealed as a result of our high sequencing depth (Supplementary Fig. 7). We noticed that at bin sizes less than 100 kilobases (kb), highly self-interacting regions begin to emerge (Fig. 1a and Supplementary Fig. 7, seen as ‘triangles’ on the heat map). These regions, which we term topological domains, are bounded by narrow segments where the chromatin interactions appear to end abruptly. We hypothesized that these abrupt transitions may represent boundary regions in the genome that separate topological domains.

To identify systematically all such topological domains in the genome, we devised a simple statistic termed the directionality index to quantify the degree of upstream or downstream interaction bias for a genomic region, which varies considerably at the periphery of the topological domains (Fig. 1b; see Supplementary Methods for details). The directionality index was reproducible (Supplementary Table 2) and pervasive, with 52% of the genome having a directionality index that was not expected by random chance (Fig. 1c, false discovery rate = 1%). We then used a Hidden Markov model (HMM) based on the directionality index to identify biased ‘states’ and therefore infer the locations of topological domains in the genome (Fig. 1a; see Supplementary Methods for details). The domains defined by HMM were reproducible between replicates (Supplementary Fig. 8). Therefore, we combined the data from the HindIII replicates and identified 2,200 topological domains in mouse ES cells with a median size of 880 kb that occupy ~91% of the genome (Supplementary Fig. 9). As expected, the frequency of intra-domain interactions is higher than inter-domain interactions (Fig. 1d, e). Similarly, FISH probes⁶ in the same topological domain (Fig. 1f) are closer in nuclear space than probes in different topological domains (Fig. 1g), despite similar genomic distances between probe pairs (Fig. 1h, i). These findings are best explained by a model of the organization of genomic DNA into spatial modules linked by short chromatin segments. We define the genomic regions between topological domains as either ‘topological boundary regions’ or ‘unorganized chromatin’, depending on their sizes (Supplementary Fig. 9).

We next investigated the relationship between the topological domains and the transcriptional control process. The *Hoxa* locus is separated into two compartments by an experimentally validated insulator^{4,7,8}, which we observed corresponds to a topological domain boundary in both mouse (Fig. 1a) and human (Fig. 2a). Therefore, we hypothesized that the boundaries of the topological domains might correspond to insulator or barrier elements.

Many known insulator or barrier elements are bound by the zinc-finger-containing protein CTCF (refs 9–11). We see a strong enrichment of CTCF at the topological boundary regions (Fig. 2b and Supplementary Fig. 10), indicating that topological boundary regions

¹Ludwig Institute for Cancer Research, 9500 Gilman Drive, La Jolla, California 92093, USA. ²Medical Scientist Training Program, University of California, San Diego, La Jolla, California 92093, USA.

³Biomedical Sciences Graduate Program, University of California, San Diego, La Jolla, California 92093, USA. ⁴Bioinformatics and Systems Biology Graduate Program, University of California, San Diego, La Jolla, California 92093, USA. ⁵Department of Statistics, Harvard University, 1 Oxford Street, Cambridge, Massachusetts 02138, USA. ⁶University of California, San Diego School of Medicine, Department of Cellular and Molecular Medicine, Institute of Genomic Medicine, UCSD Moores Cancer Center, 9500 Gilman Drive, La Jolla, California 92093, USA.

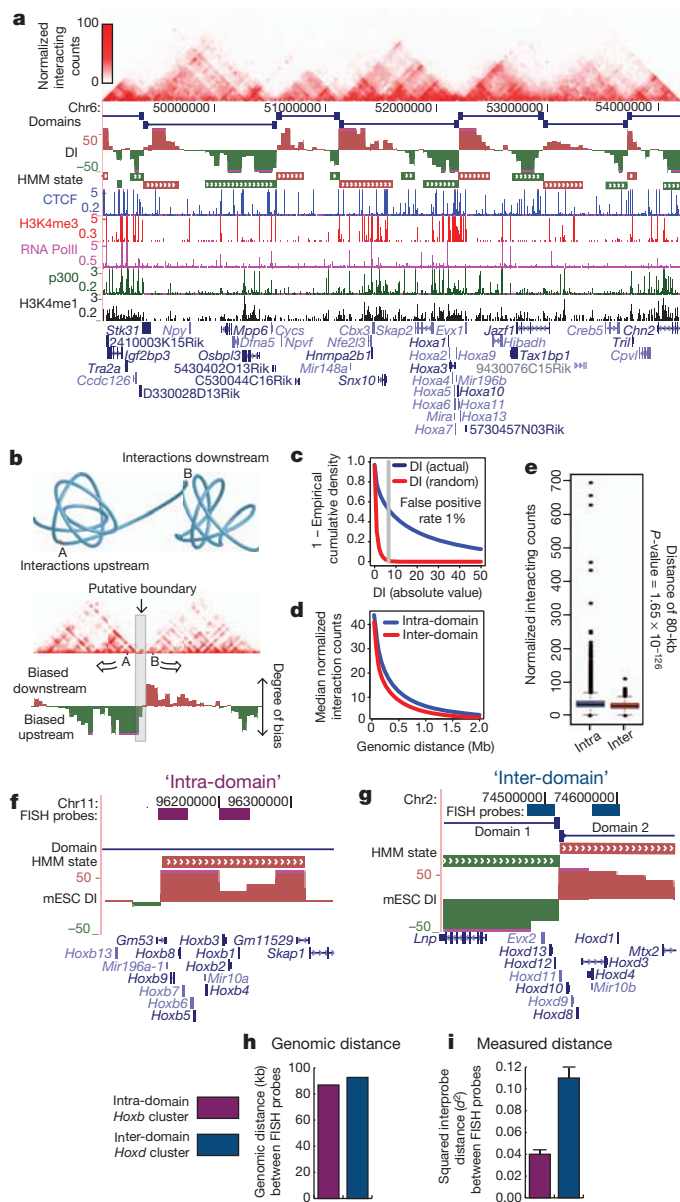


Figure 1 | Topological domains in the mouse ES cell genome. **a**, Normalized Hi-C interaction frequencies displayed as a two-dimensional heat map overlaid on ChIP-seq data (from Y. Shen *et al.*, manuscript in preparation), directionality index (DI), HMM bias state calls, and domains. For both directionality index and HMM state calls, downstream bias (red) and upstream bias (green) are indicated. **b**, Schematic illustrating topological domains and resulting directional bias. **c**, Distribution of the directionality index (absolute value, in blue) compared to random (red). **d**, Mean interaction frequencies at all genomic distances between 40 kb to 2 Mb. Above 40 kb, the intra- versus inter-domain interaction frequencies are significantly different ($P < 0.005$, Wilcoxon test). **e**, Box plot of all interaction frequencies at 80-kb distance. Intra-domain interactions are enriched for high-frequency interactions. **f-i**, Diagram of intra-domain (f) and inter-domain FISH probes (g) and the genomic distance between pairs (h). **i**, Bar chart of the squared inter-probe distance (from ref. 6) FISH probe pairs. mESC, mouse ES cell. Error bars indicate standard error ($n = 100$ for each probe pair).

share this feature of classical insulators. A classical boundary element is also known to stop the spread of heterochromatin. Therefore, we examined the distribution of the heterochromatin mark H3K9me3 in humans and mice in relation to the topological domains^{12,13}. Indeed, we observe a clear segregation of H3K9me3 at the boundary regions that occurs predominately in differentiated cells (Fig. 2d, e and Supplementary Fig. 11). As the boundaries that we analysed in

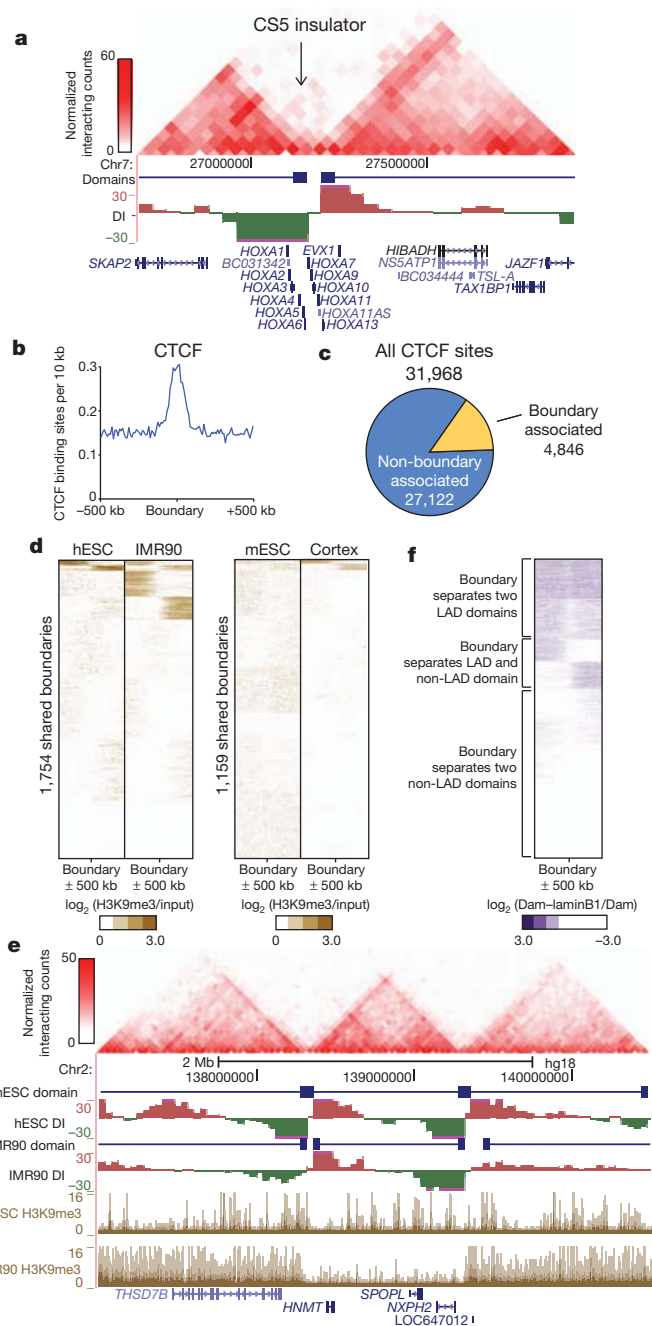


Figure 2 | Topological boundaries demonstrate classical insulator or barrier element features. **a**, Two-dimensional heat map surrounding the *Hoxa* locus and CS5 insulator in IMR90 cells. **b**, Enrichment of CTCF at boundary regions. **c**, The portion of CTCF binding sites that are considered ‘associated’ with a boundary (within ± 20 -kb window is used as the expected uncertainty due to 40-kb binning). **d**, Heat maps of H3K9me3 at boundary sites in human and mouse. **e**, UCSC Genome Browser shot showing heterochromatin spreading in the human ES cells (hESC) and IMR90 cells. The two-dimensional heat map shows the interaction frequency in human ES cells. **f**, Heat map of LADs (from ref. 14) surrounding the boundary regions. Scale is the log₂ ratio of DNA adenosome methylation (Dam)–lamin B1 fusion over Dam alone (Dam–laminB1/Dam).

Fig. 2d are present in both pluripotent cells and their differentiated progeny, the topological domains and boundaries appear to pre-mark the end points of heterochromatic spreading. Therefore, the domains do not seem to be a consequence of the formation of heterochromatin. Taken together, the above observations strongly suggest that the topological domain boundaries correlate with regions of the genome displaying classical insulator and barrier element activity, thus revealing a

potential link between the topological domains and transcriptional control in the mammalian genome.

We compared the topological domains with previously described domain-like organizations of the genome, specifically with the A and B compartments described by ref. 2, with lamina-associated domains (LADs)^{10,14}, replication time zones^{15,16}, and large organized chromatin K9 modification (LOCK) domains¹⁷. In all cases, we can see that topological domains are related to, but independent from, each of these previously described domain-like structures (Supplementary Figs 12–15). Notably, a subset of the domain boundaries we identify appear to mark the transition between either LAD and non-LAD regions of the genome (Fig. 2f and Supplementary Fig. 12), the A and B compartments (Supplementary Fig. 13, 14), and early and late replicating chromatin (Supplementary Fig. 14). Lastly, we can also confirm the previously reported similarities between the A and B compartments and early and late replication time zone (Supplementary Fig. 16)¹⁶.

We next compared the locations of topological boundaries identified in both replicates of mouse ES cells and cortex, or between both replicates of human ES cells and IMR90 cells. In both human and mouse, most of the boundary regions are shared between cell types (Fig. 3a and Supplementary Fig. 17a), suggesting that the overall domain structure between cell types is largely unchanged. At the boundaries called in only one cell type, we noticed that trend of upstream and downstream bias in the directionality index is still readily apparent and highly reproducible between replicates (Supplementary Fig. 17b, c). We cannot determine if the differences in domain calls between cell types is due to noise in the data or to biological phenomena, such as a change in the strength of the boundary region between cell types¹⁸. Regardless, these results indicate that the domain boundaries are largely invariant between cell types. Lastly, only a small fraction of the boundaries show clear differences between two cell types, suggesting that a relatively rare subset of boundaries may actually differ between cell types (Supplementary Fig. 18).

The stability of the domains between cell types is surprising given previous evidence showing cell-type-specific chromatin interactions and conformations^{5,7}. To reconcile these results, we identified cell-type-specific chromatin interactions between mouse ES cell and mouse cortex. We identified 9,888 dynamic interacting regions in the mouse genome based on 20-kb binning using a binomial test with an empirical false discovery rate of <1% based on random permutation of the replicate data. These dynamic interacting regions are enriched for differentially expressed genes (Fig. 3b–d, Supplementary Fig. 19 and Supplementary Table 5). In fact, 20% of all genes that undergo a four-fold change in gene expression are found at dynamic interacting loci. This is probably an underestimate, because by binning the genome at 20 kb, any dynamic regulatory interaction less than 20 kb will be missed. Lastly, >96% of dynamic interacting regions occur in the same domain (Fig. 3e). Therefore, we favour a model where the domain organization is stable between cell types, but the regions within each domain may be dynamic, potentially taking part in cell-type-specific regulatory events.

The stability of the domains between cell types prompted us to investigate if the domain structure is also conserved across evolution. To address this, we compared the domain boundaries between mouse ES cells and human ES cells using the UCSC liftover tool. Most of the boundaries appear to be shared across evolution (53.8% of human boundaries are boundaries in mouse and 75.9% of mouse boundaries are boundaries in humans, compared to 21.0% and 29.0% at random, P value $< 2.2 \times 10^{-16}$, Fisher's exact test; Fig. 3f). The syntenic regions in mouse and human in particular share a high degree of similarity in their higher order chromatin structure (Fig. 3g, h), indicating that there is conservation of genomic structure beyond the primary sequence of DNA.

We explored what factors may contribute to the formation of topological boundary regions in the genome. Although most topological boundaries are enriched for the binding of CTCF, only 15% of CTCF

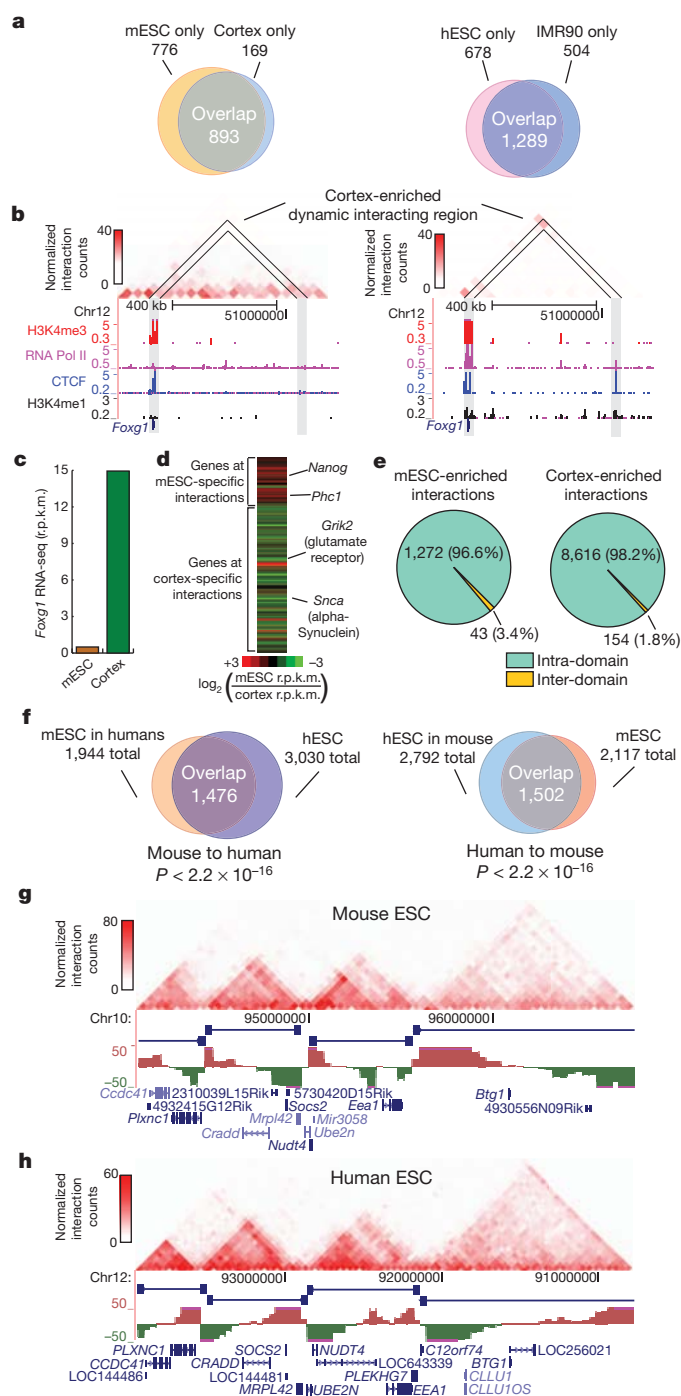


Figure 3 | Boundaries are shared across cell types and conserved in evolution. **a**, Overlap of boundaries between cell types. **b**, Genome browser shot of a cortex enriched dynamic interacting region that overlaps with the *Foxg1* gene. **c**, *Foxg1* expression in reads per kilobase per million reads sequenced (r.p.k.m.) in mouse ES cells and cortex as measured by RNA-seq. **d**, Heat map of the gene expression ratio between mouse ES cell and cortex of genes at dynamic interactions. **e**, Pie chart of inter- and intra-domain dynamic interactions. **f**, Overlap of boundaries between syntenic mouse and human sequences ($P < 2.2 \times 10^{-16}$ compared to random, Fisher's exact test). **g, h**, Genome browser shots showing domain structure over a syntenic region in the mouse (**g**) and human (**h**) ES cells. Note: the region in humans has been inverted from its normal UCSC coordinates for proper display purposes.

binding sites are located within boundary regions (Fig. 2c). Thus, CTCF binding alone is insufficient to demarcate domain boundaries. We reasoned that additional factors might be associated with topological boundary regions. By examining the enrichment of a variety of

histone modifications, chromatin binding proteins and transcription factors around topological boundary regions in mouse ES cells, we observed that factors associated with active promoters and gene bodies are enriched at boundaries in both mouse and humans (Fig. 4a and Supplementary Figs 20–23)^{19,20}. In contrast, non-promoter-associated marks, such as H3K4me1 (associated with enhancers) and H3K9me3, were not enriched or were specifically depleted at boundary regions (Fig. 4a). Furthermore, transcription start sites (TSS) and global run on sequencing (GRO-seq)²¹ signal were also enriched around topological boundaries (Fig. 4a). We found that housekeeping genes were particularly strongly enriched near topological boundary regions (Fig. 4b–d; see Supplementary Table 7 for complete GO terms enrichment). Additionally, the tRNA genes, which have the potential to function as boundary elements^{22,23}, are also enriched at boundaries (P value < 0.05 , Fisher's exact test; Fig. 4b). These results suggest that high levels of transcription activity may also contribute to boundary formation. In support of this, we can see examples of dynamic changes in H3K4me3 at or near some cell-type-specific boundaries that are cell-type specific (Supplementary Fig. 24). Indeed, boundaries associated with both CTCF and a housekeeping gene account for nearly one-third of all topological boundaries in the genome (Fig. 4e and Supplementary Fig. 24).

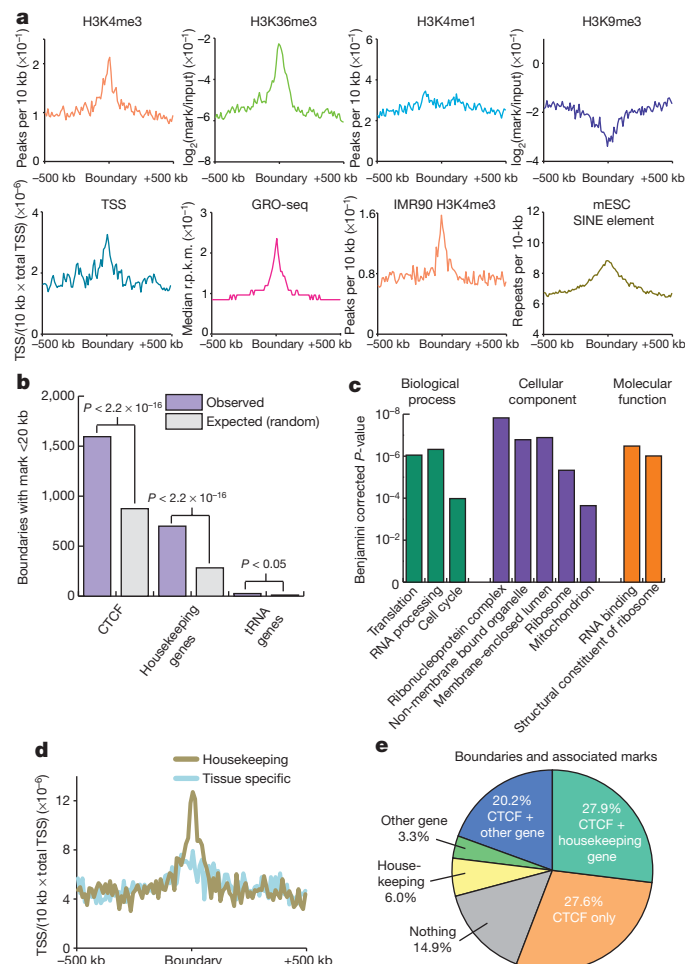


Figure 4 | Boundary regions are enriched for housekeeping genes.

a, Chromatin modifications, TSS, GRO-seq and SINE elements surrounding boundary regions in mouse ES cells or IMR90 cells. **b**, Boundaries associated with a CTCF binding site, housekeeping gene, or tRNA gene (purple) compared to expected at random (grey). **c**, Gene Ontology P -value chart. **d**, Enrichment of housekeeping genes (gold) and tissue-specific genes (blue) as defined by Shannon entropy scores near boundaries normalized for the number of genes in each class (TSS/10 kb/total TSS). **e**, Percentage of boundaries with a given mark within 20 kb of the boundaries.

Finally, we analysed the enrichment of repeat classes around boundary elements. We observed that Alu/B1 and B2 SINE retrotransposons in mouse and Alu SINE elements in humans are enriched at boundary regions (Fig. 4a and Supplementary Figs 24 and 25). In light of recent reports indicating that a SINE B2 element functions as a boundary in mice²⁴, and SINE element retrotransposition may alter CTCF binding sites during evolution²⁵, we believe that this contributes to a growing body of evidence indicating a role for SINE elements in the organization of the genome.

In summary, we show that the mammalian chromosomes are segmented into megabase-sized topological domains, consistent with some previous models of the higher order chromatin structure^{1,26,27}. Such spatial organization seems to be a general property of the genome: it is pervasive throughout the genome, stable across different cell types and highly conserved between mice and humans.

We have identified multiple factors that are associated with the boundary regions separating topological domains, including the insulator binding factor CTCF, housekeeping genes and SINE elements. The association of housekeeping genes with boundary regions extends previous studies in yeast and insects and suggests that non-CTCF factors may also be involved in insulator/barrier functions in mammalian cells²⁸.

The topological domains we identified are well conserved between mice and humans. This indicates that the sequence elements and mechanisms that are responsible for establishing higher order structures in the genome may be relatively ancient in evolution. A similar partitioning of the genome into physical domains has also been observed in *Drosophila* embryos²⁹ and in high-resolution studies of the X-inactivation centre in mice (termed topologically associated domains or TADs)³⁰, indicating that topological domains may be a fundamental organizing principle of metazoan genomes.

METHODS SUMMARY

Cell culture and Hi-C experiments. J1 mouse ES cells were grown on gamma-irradiated mouse embryonic fibroblasts cells under standard conditions (85% high glucose DMEM, 15% HyClone FBS, 0.1 mM non-essential amino acids, 0.1 mM β -mercaptoethanol, 1 mM glutamine, LIF 500 U ml⁻¹, 1 \times Gibco penicillin/streptomycin). Before collecting for Hi-C, J1 mouse ES cells were passaged onto feeder free 0.2% gelatin-coated plates for at least two passages to rid the culture of feeder cells. H1 human ES cells and IMR90 fibroblasts were grown as previously described¹³. Collecting the cells for Hi-C was performed as previously described, with the only modification being that the adherent cell cultures were dissociated with trypsin before fixation.

Sequencing and mapping of data. Hi-C analysis and paired-end libraries were prepared as previously described² and sequenced on the Illumina Hi-Seq2000 platform. Reads were mapped to reference human (hg18) or mouse genomes (mm9), and non-mapping reads and PCR duplicates were removed. Two-dimensional heat maps were generated as previously described².

Data analysis. For detailed descriptions of the data analysis, including descriptions of the directionality index, hidden Markov models, dynamic interactions identification, and boundary overlap between cells and across species, see Supplementary Methods.

Received 26 September 2011; accepted 27 March 2012.

Published online 11 April 2012.

- Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889 (2010).
- Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- Yaffe, E. & Tanay, A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nature Genet.* **43**, 1059–1065 (2011).
- Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011).
- Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
- Eskeland, R. *et al.* Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol. Cell* **38**, 452–464 (2010).
- Noordermeer, D. *et al.* The dynamic architecture of Hox gene clusters. *Science* **334**, 222–225 (2011).

8. Kim, Y. J., Cecchini, K. R. & Kim, T. H. Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc. Natl Acad. Sci. USA* **108**, 7391–7396 (2011).
9. Phillips, J. E. & Corces, V. G. CTCF: master weaver of the genome. *Cell* **137**, 1194–1211 (2009).
10. Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).
11. Handoko, L. *et al.* CTCF-mediated functional chromatin interactome in pluripotent cells. *Nature Genet.* **43**, 630–638 (2011).
12. Xie, W. *et al.* Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell* **148**, 816–831 (2012).
13. Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479–491 (2010).
14. Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
15. Hiratani, I. *et al.* Genome-wide dynamics of replication timing revealed by *in vitro* models of mouse embryogenesis. *Genome Res.* **20**, 155–169 (2010).
16. Ryba, T. *et al.* Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**, 761–770 (2010).
17. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature Genet.* **41**, 246–250 (2009).
18. Scott, K. C., Taubman, A. D. & Geyer, P. K. Enhancer blocking by the *Drosophila* gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* **153**, 787–798 (1999).
19. Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* **23**, 2484–2489 (2009).
20. Marson, A. *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521–533 (2008).
21. Min, I. M. *et al.* Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* **25**, 742–754 (2011).
22. Donze, D. & Kamakaka, R. T. RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in *Saccharomyces cerevisiae*. *EMBO J.* **20**, 520–531 (2001).
23. Ebersole, T. *et al.* tRNA genes protect a reporter gene from epigenetic silencing in mouse cells. *Cell Cycle* **10**, 2779–2791 (2011).
24. Lunyak, V. V. *et al.* Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* **317**, 248–251 (2007).
25. Schmidt, D. *et al.* Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **148**, 335–348 (2012).
26. Jhunjhunwala, S. *et al.* The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell* **133**, 265–279 (2008).
27. Capelson, M. & Corces, V. G. Boundary elements and nuclear organization. *Biol. Cell* **96**, 617–629 (2004).
28. Amouyal, M. Gene insulation. Part I: natural strategies in yeast and *Drosophila*. *Biochem. Cell Biol.* **88**, 875–884 (2010).
29. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
30. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* <http://dx.doi.org/10.1038/nature11049> (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful for the comments from and discussions with Z. Qin, A. Desai and members of the Ren laboratory during the course of the study. We also thank W. Bickmore and R. Eskeland for sharing the FISH data generated in mouse ES cells. This work was supported by funding from the Ludwig Institute for Cancer Research, California Institute for Regenerative Medicine (CIRM, RN2-00905-1) (to B.R.) and NIH (B.R. R01GH003991). J.R.D. is funded by a pre-doctoral training grant from CIRM. Y.S. is supported by a postdoctoral fellowship from the Rett Syndrome Research Foundation.

Author Contributions J.R.D. and B.R. designed the studies. J.R.D., A.K., Y.L. and Y.S. conducted the Hi-C experiments; J.R.D., S.S. and F.Y. carried out the data analysis; J.S.L. and M.H. provided insight for analysis; F.Y. built the supporting website; J.R.D. and B.R. prepared the manuscript.

Author Information All Hi-C data described in this study have been deposited in the GEO under accession number GSE35156. We have developed a web-based Java tool to visualize the high-resolution Hi-C data at a genomic region of interest that is available at <http://chromosome.sdsc.edu/mouse/hi-c/>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.R. (biren@ucsd.edu).